
A SURVEY ON CHURN ANALYSIS IN VARIOUS BUSINESS DOMAINS

A PREPRINT

Jaehuyn Ahn
Intelligence Labs
Nexon Korea
Gyeonggi-do, Republic of Korea
jaehyunahn@nexon.co.kr

Junsik Hwang
Intelligence Labs
Nexon Korea
Gyeonggi-do, Republic of Korea
junsik.hwang@nexon.co.kr

Doyoung Kim
Intelligence Labs
Nexon Korea
Gyeonggi-do, Republic of Korea
ehdud8565@nexon.co.kr

HyukGeun Choi
Intelligence Labs
Nexon Korea
Gyeonggi-do, Republic of Korea
hgchoi16@nexon.co.kr

Shinjin Kang
School of Games
Hongik University
Sejong Campus, Republic of Korea
directx@hongik.ac.kr

October 27, 2020

ABSTRACT

In this paper, we 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. In this paper, we collected the definitions of churn used in the fields of business administration, marketing, IT, telecommunications, newspapers, insurance and psychology, and described their differences. Based on this, we classified and explained churn loss, feature engineering, and prediction models. 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 Relationship 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 fulfil 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, we 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 Chapters 4 and 5 introduce losses and features used in Churn, respectively. Chapter 6 introduces typical churn-based prediction models by classifying them according to each algorithm, and presents which algorithm is mainly used in each industry.

2 Definition of Churn

Churn has been defined in various ways in multiple industries. In this chapter, we 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

Table 1: Customer churn decision by period and method

Setting	Churn observation criteria	Publishing Information
Contractual	Monthly	Chen, Yian, et al. (2018) [15]
Contractual	Monthly	Mozer, Michael C., et al. (2000) [16]
Contractual	Monthly	Dahiya, Kiran, and Surbhi Bhatia. (2015) [65]
Contractual	Monthly	Bahnsen, Alejandro Correa, Djamila Aouada, and Björn Ottersten. (2015) [66]
Contractual	Monthly	Coussement, Kristof, and Dirk Van den Poel. (2008) [67]
Contractual	Monthly	Radosavljevik, Dejan, Peter van der Putten, and Kim Kyllèsbech Larsen. (2010) [68]
Contractual	Monthly	Glady, 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 del 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	Ahn et al. (2006) [33]
Contractual	Monthly	Wei, 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	Athanassopoulos (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.

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, we can 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 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.

	Observation Period (Log data)									
	Before window			Churn Determination Window				After window		
	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8	Week 9	Week 10
User A	●	●	●				Churned			
User B	●	●	●				Churned	●	●	●
User C	●		●				Churned			●
User D	●	●	●				Survived			
User E	●	●		●	●	●	Survived	●	●	
User F	●	●	●				Survived	●		

Figure 1: The result of schematic representation of the time window method used for churn log analysis in non-contractual case. According to this figure, data was collected for 10 weeks and the time window is set from Week 4 to Week 7.

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] [141]. 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].

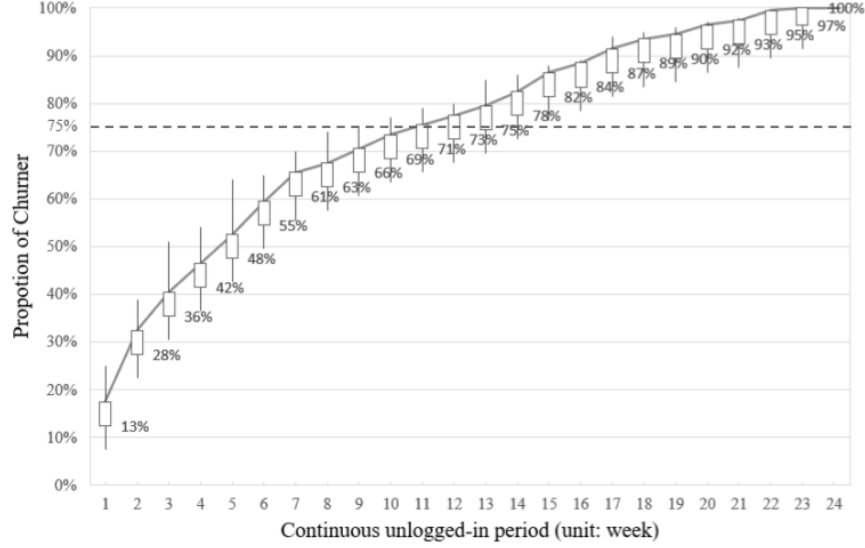


Figure 2: In order to determine the time window period, researchers should collect customers' continuous unconnected periods and draw a cumulative graph. According to the figure, the period during which 75% or more customers are continuously unconnected is 14 weeks. Therefore, if a customer is unconnected for 14 weeks, he/she is regarded as churn and this period is a time window section.

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 cases of churn prediction for each business field are summarized in Appendix A.

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]. Larivière, 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. Gladly, 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 Customer Churn Loss

Customer churn behavior is quantitative. However, it is difficult to directly relate customer churn to a decrease in sales. Therefore, of the studies on churn prediction, there is a study that introduces a method of calculating the loss of a single customer. In this way, we can calculate the value of a churn prevention model by multiplying the loss cost of one customer by the number of people who are prevented from churn with a churn prediction algorithm.

4.1 Customer Acquisition Cost (CAC)

The customer acquisition cost (CAC) is the total cost that is spent until a customer is convinced of a service. CAC can be calculated by simply dividing all cost spent on acquiring customers, marketing campaign for example, by the number of customers acquired in the period the money was spent. The company measures the cost of acquiring a customer with the CAC, and that CAC cost is the minimum value to operate service makes from a customer with Return on Investing (ROI). The CAC occurs mainly through marketing activities. If a customer churns from a service, the company will have to recruit another customer by spending the CAC to maintain the service. According to the study conducted by Mozer, Michael C., et al. (2000), the retained customers are known to provide a better return on investment than the newly recruited customers through the CAC [16]. In such way, if a customer who is likely to churn the service in the near future can be inferred through the churn prediction, the basis for measuring a suitable incentive for the customer while minimizing the CAC can be established. In addition, by multiplying the number of customers planning to churn by the incentive cost to be provided to those customers, the cost of business loss that can incur when the customers are not retained can be calculated through the churn prediction model [45] [53]. Therefore, some studies measured the CAC, and calculated it as the loss incurred when a customer churned [8] [31] [72].

4.2 Customer Lifetime Value (CLV)

The customer lifetime value (CLV) is the cost that a customer expects to pay when acquiring a customer. This reason why this cost is important is that the CLV is the expected earnings from the customer's use of a service when acquiring a new customer, and the CLV is a useful indicator for setting the upper bound when calculating customer-related costs. Marketing costs and incentives provided for customers who are going to churn are typical examples of customer-related costs [52]. The efficiency of a retained customer value is usually calculated with the CLV. A retained customer represents a customer who the churn prediction model predicted to churn in the near future but survived after receiving

incentives. This is because the cost can vary depending on the company's policy and marketing timing for the CAC method. There are multiple studies on the method of calculating the CLV and that apply the CLV to churn models. Verbraken, Thomas, Wouter Verbeke, and Bart Baesens. (2012) and Neslin, Scott A., et al. (2006) proposed formulas for calculating the net profit using churn rate, CAC, CLV, fixed operating cost, and incentive cost [54] [55]. Additionally, the same approach has been taken by Fader, Peter S., and Bruce GS Hardie. (2009) as well [149]. In the formula for deriving CLV, the survival rate (retention) γ for the customer's time period t should be derived first. γ can be derived through the probabilistic distribution as well. When γ denotes a retention rate, $1 - \gamma$ expresses a churn rate. In this case, the expected survival time of a customer can be expressed as $\frac{1}{1-\gamma}$. Lastly, assuming that the profit contribution cost per customer (customer value) for t period is expressed as m , the CLV can be obtained by $m\gamma^t$. Here, the value of m may be different for each customer segment. In the case of calculating m value of new acquisition customer, m is derived by dividing the net profit from active customers for time t by the number of active customers. Further, the contribution of a specific segment can be calculated by dividing the net profit from activated specific segment customers for t period by the number of activated segment customers during t period. In the case where the customer value is discounted during the time t , the discount rate is defined as d , and the discount value during the time t is expressed as $\frac{1}{(1+d)^t}$. Ultimately, the CLV having discount term for a given time t can be expressed as $\frac{m\gamma^t}{(1+d)^t}$. Since the CLV should include the concept of lifetime, time t can be generalized as follows.

$$CLV = \sum_{t=0}^T m \frac{\gamma^t}{(1+d)^t}$$

As a way of calculating the loss of employee churn, Saradhi, V. Vijaya, and Girish Keshav Palshikar. (2011) calculated churn rates, and by using the method of multiplying customer value by the remaining survival time, the authors calculated the projected value in future time of churned employees who failed to fulfill the CLV [50]. Based on the formula for deriving net cashflow using survival time parameters suggested by Reinartz, Werner J., and Vijay Kumar. (2000), the study conducted by Gladys, Nicolas, Bart Baesens, and Christophe Croux. (2009) used the approach of multiplying individual cashflows for entire product to calculate the CLV in the retail service field [38] [146].

5 Feature Engineering

Churn is generally related to customers' last time activity. However, predicting churn and compensating for it with the last log of service usage do not change the overall service usage patterns of customers. Therefore, some studies show that short-term prediction and monetary rewards soon leads to another churn [10] [45]. Therefore, studies have emerged in recent years to develop other features that are as important as the last log the customer left before churning, or to discover potential churners by reprocessing the time series features.

5.1 Developing New Feature

Sifa, Rafet, Christian Bauckhage, and Anders Drachen. (2014) diagrammed the related signs leading to churn, and grouped the features corresponding to each diagram and managed them in Game field [56]. This study focused on detecting signs that led to churn rather than building churn models and comparing performance. They linked unmeasurable numbers such as key indicators of services, the number of complaints raised and psychological fluctuations to service features so as to measure them, and utilized them for the research on churn prediction.

Yang, Wanshan, et al. (2019) judged that the probability of churn would increase when the regularity of customers' behavior using services was broken [43]. They added the change in the customer service playtime distribution as a feature, and maintained that the feature was of a great help in estimating churn by creating a machine learning model. Hadiji, Fabian, et al. (2014) and Yang, Wanshan, et al. (2019) predicted churn using KPI features [42] [43]. They used the indicators used in business administration as churn prediction features since churn was related to management indicators.

Runge, Julian, et al. (2014) associates the value of service goods possessed by customers with customer churn. This study intensively used features related to assets (e.g., reserves, items) in customers' services [45]. Paid goods, free goods, last purchase, and last purchase date, and so on were used as assets. They assumed that as the user had more goods to use in the service, the opportunity cost became larger, which would be a major indicator of churn.

In the finance field, Chu, Tsai, and Ho (2007) used customer demographic CRM features and business branch relationships to conduct churn prediction [137]. They have predicted customer churn using the customer information such as gender, zip code, and customer's industry code and the service provider information such as tenure, time of service suspended, and average invoice. In order to predict customer churn of Pay-TV services, Burez, Jonathan, and

Dick Van den Poel (2008) selected customer behavioral loyalty features and combined them with CRM features [132]. In addition to using service information such as payment type and contract expiration month, the authors used demographic CRM features including customer's age, province, and customer type and additional features for classifying disloyal customers- namely bad payment behaviors, number of notices to pay, and number of deactivation of the device- in order to conduct churn prediction.

Logs remain in the majority of services. However, not all logs are helpful in estimating churn. Mozer, Michael C., et al. (2000) maintained that in general the indicators that measured the quality of services were good data for estimating churn in telecommunication field [16]. Dror, Gideon, et al. (2012), by collecting responses such as like and dislike from the Internet service they ran, used them to predict churn. They explained that customer satisfaction was expressed as an emotional expression, which was a direct expression of service satisfaction [57].

Fader, Peter S., Bruce GS Hardie, and Ka Lok Lee. (2005) and Glady, Nicolas, Bart Baesens, and Christophe Croux. (2009) applied the RFM method, which is used for selecting loyal customers in the marketing field, to the churn prediction [38] [49]. The RFM stands for recency (latest service transaction), frequency (service transaction frequency) and, monetary (customer's purchase size). These features are used to extract customers who carry business values. In terms of marketing, the RFM features are generally used to classify customers into five groups based on the RFM scores [140]. However, in the above two studies, the authors characterized RFM features as the important features that can derive the net business values of customer-service relationship and conducted churn prediction based on the premise that loyal customers have higher service stickiness and switching cost. On the other hand, to conduct a customer churn prediction, Tamaddon Jahromi, Ali, et al. (2010) applied RFM features to the telecommunication field and added about 12 new features, including the latest telecommunication service subscription period, call frequency, total call cost, etc. [20]. The authors also used mobile carrier-specific features including call time, number of incoming or outgoing calls, and total talk time between specific customers to conduct churn prediction. Further, Wei, Chih-Ping, and I-Tang Chiu. (2002) have utilized RFM features in the telecommunication field as well [41]. By using the time length between the contract starting date and termination date, the authors set frequency of service use as the recency feature, payment type as the frequency feature, and payment type as the monetary feature. Additionally, as for the mobile carrier-specific feature, the authors derived an influence feature, which indicates the number of distinct receivers the customer called in the outgoing call list. Buckinx, Wouter, and Dirk Van den Poel. (2005) applied RFM features to the retail field to conduct churn prediction [96]. The authors used features of customer's recent purchase or consumption time of the day (Recency), number of purchases (Frequency), and the amount of spending (Monetary). As for the retail service-specific features, metadata such as customer-supplier relationship, buying categories, mode of payment, brand purchase behavior, and usage of promotions were used.

5.2 Feature modification

Pure log data includes all truthful customer information. However, churn prediction tends to be more accurate when processing raw logs. This is because service indicators are generally sparse and have many outliers, and the data distribution is skewed to one side. Therefore, it is necessary to extract important information using an appropriate feature engineering technique for churn prediction.

In general, there have been few papers mentioning feature engineering know-how about churn. However, Zhang, Rong, et al. (2017) shared useful information when building an algorithm to predict churn from log data [11].

5.2.1 One-hot encoding

One-hot encoding is primarily a feature engineering method for nominal categorical data. In order to implement machine learning on categorical data without using a tree-based algorithm, the data must be converted to numerical data. The algorithm used in this instance is called the one-hot encoding. One-hot encoding is expressed by a group of bits among which the legal combinations of values are only those with a single high (1) bit and all the others low (0) bits. One-hot encoding on categorical data in this way produces an orthogonal feature space for each category. Although alternative methods, such as numerical encoding or binary encoding, exist for processing categorical data, each method has their drawbacks: Numerical encoding breaks the nominal nature inherent in categorical data and results in encoding with linearity between categories, likewise binary encoding produces distance between categories. These changes have the side effect of the model learning unintentional continuity. Therefore, categorical data should be modified into one-hot encoding features.

5.2.2 Bucketing

Outliers are problematic for service data. Bucketing (also called Discrete binning or Data binning) can be used for both categorical and continuous features. For features that are too sparse to be used in the model due to the large variance

because the feature value category is too wide, bucketing is a technique that makes these features into categorical features. For example, a positive integer feature x can be modified into three discrete bucket: 1 (if $0 \leq x < 5$), 2 ($5 \leq x < 10$), and 3 (if $10 \leq x$).

5.2.3 Data Imputation

Data imputation is the process of replacing missing data with substituted values. It is recommended to fill the missing data by any means. Nimmagadda, Sravya, Akshay Subramaniam, and Man Long Wong. (2017) argued that filling the missing data improved performance instead of dropping them [58]. Sifa, Rafet, et al. (2015) improved performance by using a semi-supervised learning technique since there was no sufficient data to solve prediction problems [52]. Data imputation techniques may help to solve data imbalances and make the best of the information used in the model.

5.2.4 Normalization

Normalization is data pre-processing technique for stability in several training machine learning algorithms. This process scale individual samples to have unit norm. The distribution of service data is generally skewed to one side. Without normalization, the machine learning model generally have to select a very small learning rate when searching for the optimum, resulting in a long training time. Normalized features can achieve rapid model training while using relatively large learning rates compared to the initial data. Nimmagadda, Sravya, Akshay Subramaniam, and Man Long Wong. (2017) used techniques such as log-normalization or quantile normalization to build a prediction model other than 0-1 normalization [58].

5.2.5 Feature Embedding

Schweidel, David A., and George Knox. (2013) proposed a parsimonious model by integrating customer behavior data into latent attrition models in order to provide direct marketing target selection [142]. This model enables dense embedding of sparse customer data by extracting latent attritions. In general, when the customer behavior is stored as log data in gaming, Internet service, and telecommunication service fields, the customer data is often found in the form of high-dimensional sparse data. Hence, in the past, the customer behavior data was simplified through a bucketing process. However, with the emergence of deep learning techniques of machine learning, it has become possible to manage time-dependent high-dimensional sparse data. Unlike the explicit methods such as bucketing, a deep learning algorithm can learn that customers' latent behavior who are about to churn in a end-to-end way. Moreover, a deep learning algorithm enables generating latent features by compressing long-term features that is a new technique of feature embedding. An autoencoder is one of the example techniques to this. Autoencoder is trained based on the encoding and decoding process, where latent vectors are generated during the process. The latent features generated during this process compresses high-dimensional sparse data into low-dimensional dense data. It has been suggested that the model that uses these obtained vectors as input of the fully connected networks provides better prediction performance than the traditional model that uses the sparse features as is. For example, to predict the future demand of Uber customers, Zhu, Lingxue, and Nikolay Laptev. (2017) improved the performance of the prediction model by compressing sparse data through a long-short-term memory autoencoder and then concatenating the data with fully connected neural networks (FCNNs) [143]. This trend is found in the churn prediction as well. Lee, Eunjo, et al. (2018) claimed that in the customer churn prediction competition hosted by the authors, the winning team showed significantly better performance than other teams by utilizing the autoencoder to compress the sparse data [10]. Zhang, Rong, et al. (2017) also used latent feature modification for churn prediction. They classified features into a memorization feature and a generalization feature depending on data attributes [11]. The memorization feature refers to data that is likely to be a latent feature among time series data. These types of data slowly reveal the characteristics of the churn/no-churn group usually over a period of time. This data is modified by a deep learning model LSTM, embedding, or autoencoder, and then processed by converting them into a dense-low dimension latent feature. As opposed to the memorization feature, the generalization feature exhibits the characteristics of the churn/no-churn group with only short-term attributes. This feature can represent the attributes of the churn/no-churn group with a short-term section or value itself. The generalization feature data predicts results with shallow machine learning models such as logistic linear regression. The churn feature that combines the memorization feature and the generalization feature predicts churn by combining deep learning models and traditional machine learning techniques.

5.3 Dealing with imbalanced data

In a stable service, the number of churned customers generally account for a small proportion compared to the number of retained customers. For example, suppose that 96 percent of the data is composed of retained customers and 4 percent is churned customers. If a prediction model is trained using this data and it always outputs only the results that indicate customers being retained, the model would maintain a 96 percent accuracy. In this case, although the

accuracy of the model may be said to be high, the model would not be effective in identifying characteristics of churned customers. Thus, the most ideal method is adjusting the churned data and retained data to have a similar proportion. This is because a balanced dataset composed of churned and retained customers has higher noise tolerance than an imbalanced dataset, hence it is more likely to be able to obtain decision boundaries for minor groups, which in this case denotes the churned customer dataset [91]. Meanwhile, there could be an issue where the prediction model may not have enough data for training if a simple undersampling method is used to obtain a balanced dataset. To address this, Burez, Jonathan, and Dirk Van den Poel. (2009) utilized a method called the CUBE method to improve the churn prediction performance. [95]. In addition, Amin, Adnan, et al. (2016) obtained the IBM telecommunication dataset and applied the oversampling method on the imbalanced dataset. Subsequently, the authors used the generated balanced dataset to train a churn prediction model [144]. The authors also compared and analyzed the performance of the churn prediction models that implemented widely used oversampling methods. According to the study, the Mega-Trend Diffusion Function (MTDF) method provided the highest accuracy when compared with other techniques such as Synthetic Minority Oversampling Technique (SMOTE), Adaptive Synthetic Sampling Approach (ADASYN), Majority Weighted Minority Oversampling Technique (MWMOTE), Immune Centroids Oversampling Technique (ICOTE) and couples Top-N Reverse K-Nearest Neighbor (TRkNN) algorithm. Gui, Chun. (2017) applied undersampling, oversampling, and SMOTE methods to an imbalanced dataset of telecommunication field and compared the performance of the derived churn prediction models [145]. In the study, the author suggested that the SMOTE sampling technique provided the best prediction performance. The use of undersampling technique when dealing with churn dataset has disclosed the possibility of discarding useful information that expresses the characteristics of retained customers. On the contrary, the use of oversampling technique has disclosed the possibility of raising the overfitting issue due to replicating insufficient variance size of the oversampled churned group data.

6 Churn Prediction Models

6.1 Building Churn Prediction Models

There are four domains as a method of building a churn prediction model. They are traditional machine learning, statistics, graph theory, and deep learning. In Appendix B.1, we summarized the papers that built churn prediction models based on these four criteria.

The boundary is blurring between the above four disciplines in recent times. However, Breiman, Leo. (2001) divided statistics and machine learning into a stochastic data model and an algorithmic model, which influenced the way that Matthew Stewart. (2019) and Bzdok, D., Altman, N. and Krzywinski, M. (2018) divided statistics and machine learning [60] [61] [62]. Witten, Ian H., and Eibe Frank. (2002) described that machine learning techniques had developed with data mining since the advent of computers while statistics had focused on mathematics-based hypothesis tests [63].

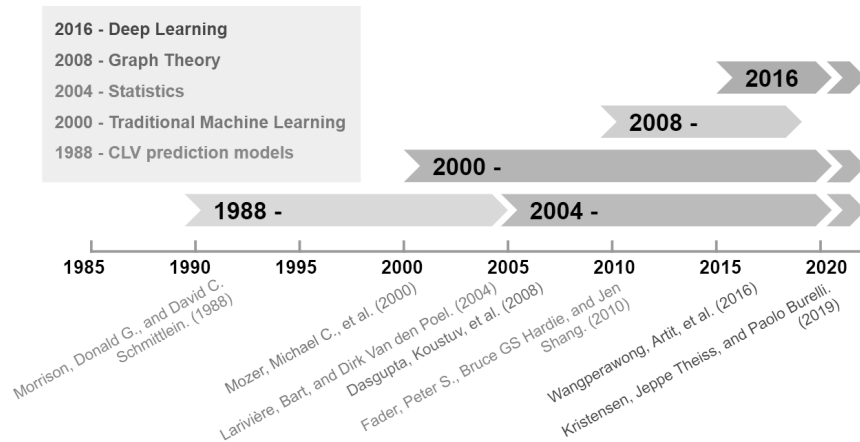


Figure 3: The algorithms used to predict customer churn are organized by year. CLV prediction algorithm is combined with calculating customer profitability and churn rate. In order to measure CLV, churn prediction model must be built first. Statistical methods such as sBG or NBD are used for predicting CLV. However, the “Statistics” depicted in this Figure defines the first emergence of accepted literature where the contribution was a model constructed via statistical methods for predicting only customer churn prediction.

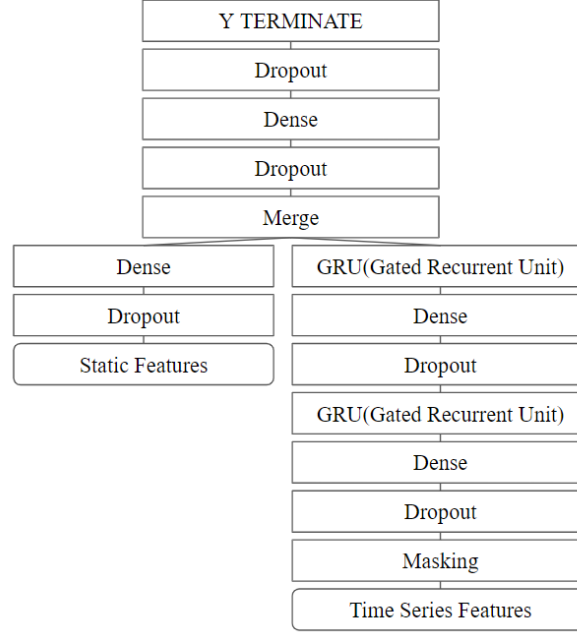


Figure 4: An example of the deep learning model structure that predicts churn by combining time series log data and static features data

In statistics, probability models have been mainly used for conducting churn predictions. In particular, probability models have been traditionally used for customer-base analysis [149]. In a customer-base analysis, churn rate is applied to the survival time estimation when calculating the CLV. Figure 3 illustrates that the appearance of churn prediction algorithms by year. CLV prediction algorithm is combined with calculating customer expected revenue and churn rate prediction model. To calculate CLV within the contractual settings, a shifted-Beta-Geometric (sBG) model is used. The sBG model uses beta distribution to make shifts for every instances of change in time t in order to fit the retention rate. Accordingly, the sBG model allows continuous interpretation of the conditions in which the customer retention is determined to be discrete-time contractual in a contractual service [150]. In the non-contractual settings, the repeat-buying behavior of customers have been previously expressed through negative binomial distributions (NBD) [151]. Further, the distribution of churn used gamma mixture of exponential, which is also known as the Pareto (of the second kind) distribution. By combining the buyer behavior and survival distribution, the CLV can be calculated. and this method is simply referred to as the Pareto/NBD model. The Pareto/NBD method has been actively used as a probability model for deriving the CLV until recently [19] [147] [148]. As another method, the beta-geometric/beta-binomial (BG/BB) model can be used. In this model, the beta-geometric model fits the retention rate and the beta-binomial distribution fits the consumer purchasing behavior [149]. In the non-contractual settings, customer churn tends to have a characteristic of continuous probabilistic. In the non-contractual settings, customer churn is not easy to define and trace. To define customer churn in the non-contractual settings, researchers used in time-series modification techniques such as grouping customer id to make tidy data or calculate behavioral variances [114]. Utilizing the features produced from this processed data, researchers begin predicting customer churn. The statistical model used here is based on statistical inference and hypothesis testing, and survival analysis with hazard methods are used to build churn prediction models. Machine learning techniques also began to be utilized in customer churn for non-contractual settings. Compared to statistical methods, machine learning techniques have robust non-linear relationships between features and can learn heterogeneous effect when given diverse features. The graph theory identifies churn as a mathematical relationship. It configures graph attributes by feature and by customer, and expresses their relationship as edges. Once a graph is built, it searches churning customers through the graph correlation analysis. Recently, deep learning techniques have emerged as a method of predicting customer churn. Deep learning is an intense extension of machine learning with neural network algorithms. However, deep learning also has many variances and tend to be classified separately from conventional machine learning. Deep learning techniques often used for predicting customer churn mainly involves training via sparse customer data that have been densified, or a fully-connected neural network produced from extracting the latent vector from the autoencoder and concatenating those features with static features.

The deep learning model is a relatively recent analysis method of predicting churn. According to Goodfellow, Ian, Yoshua Bengio, and Aaron Courville. (2016), deep learning is part of machine learning [64]. However, because

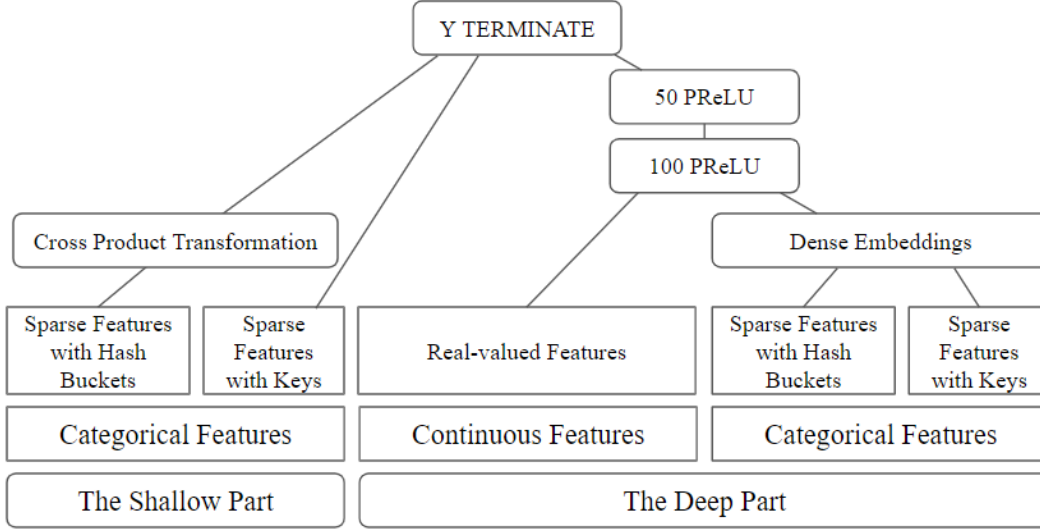


Figure 5: An example of the deep and shallow model structure to predict churn

its academic significance has recently grown, it has established itself as a single academic field. This is true for building churn prediction analysis models. Lee, Eunjo, et al. (2018) disclosed that a model using deep learning predicted customer churn with a higher probability than a traditional machine learning model for game churn prediction analysis [10]. Fig. 4 schematically represents the model of the team that won this competition. They summarized the features using the memorization and generalization techniques described in feature modification, and increased churn predictability by combining the deep learning model and the traditional machine learning model. Zhang, Rong, et al. (2017) compared the traditional machine learning model and the deep learning model in customer churn prediction problems in the insurance industry [11]. Fig. 5 is a schematic diagram of the Deep and Shallow model they built. Zhang, Rong, et al. (2017) classified features to be applied to the deep learning model and processed them, and then combined the results. In the study, they compared the deep learning-based churn prediction method they developed with the traditional machine learning-based churn prediction algorithm. The Deep and Shallow model they built showed excellent churn prediction performance compared to other models. In this study, although deep learning is part of machine learning, it is used as a new breakthrough algorithm for churn prediction problems. Appendix B.1 shows the classification of studies on churn based on churn prediction algorithms.

6.2 Churn Prediction Models in Various Business Field

Table 2 shows a summary of techniques that are classified by business field. We were able to confirm that preferred modelling techniques were different depending on the business field. Businesses with dense log data and easy access to customer information, such as the games and telecommunications industries, are applying relatively many deep learning techniques using big data, which is a fast trend. As for the financial and insurance sectors, since the log data is relatively small and the information obtained from customers does not change to a great degree, there are many statistical approaches using traditional machine learning models or survival analysis. The reason why the preferred model for each business field is different is that the types and cycles of log data used for each business are different. Apart from this, it seems to be different since the churn model that best interprets the relevant log data is applied.

6.3 Performance Evaluation

In general, the performance evaluation algorithms of machine learning model developed for churn prediction use the area under the curve (AUC) of receiver operating characteristic (ROC) curve or the lift. The ROC curve is drawn by plotting sensitivity values on the y-axis and false positive rate on the x-axis. A ROC curve is a very robust measurement criterion that measures classifiers independent of class distribution and misclassification error cost. In this way, x-axis denotes the proportion of non-churn cases that were incorrectly classified as churn, and y-axis is defined as the portion of churn cases that were classified correctly [121] [136]. Thus, the AUC close to 1 indicates that the churn prediction model accurately distinguishes difference between the characteristics of churn customers and non-churn customers [13] [45] [55] [95]. On the other hand, some churn studies have often used a top 10% decline lift performance metric [13] [67]. Lift is a

Table 2: Number of Churn Prediction models in Various Business Field

Business Field	Traditional Machine Learning	Statistics	Graph Theory	Deep Learning
Energy	1			
Finance	10	2		1
Telecommunications	26	3	3	1
Game	16	7		2
Marketing	1			
Psychology	1			
Human Resources	1			
Insurance	3			1
Commerce	1			
Internet Commerce	1			
Internet Service	1	1		1
Music Streaming Service	1			
Newspapers	4			
Retails	2			

Duplicate business fields included.

performance measure obtained by dividing baseline lift by the response for each fraction. When using top 10% decline lift as in the above reference studies, the customer list sorted in the descending order based on the prediction rate is divided into ten fractions. Subsequently, the lift values for each fraction are derived and the descending speed of the curve is observed. Additionally, top-decile lift technique is also often used as it enables allocating marketing budgets proportionally to customers who are more likely to churn as predicted by the churn prediction model [95].

7 Conclusions

In this study, we 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. Some of the papers introduced in 6. 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, we 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, we 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.

8 LIMITATIONS AND ISSUES FOR FURTHER RESEARCH

Churn studies on different fields are undoubtedly helpful in grasping the comprehensive view of churn and exploring various features to apply them to the churn models. However, as each study uses different sizes and types of features in the data, the set of studies provided in this paper has a limitation in comparing a common performance. Accordingly, although researchers may be able to discern whether if their constructed model is used widely through our study, they would not be able to determine which model is suitable and has the best performance for their study. Thus, in the future study, we intend to combine the feature engineering of fields introduced in this paper with the open churn datasets and construct various churn prediction models, including a deep learning model, and then conduct experiments on comparing and evaluating the suitability of each model.

A Churn Analysis in Various Business Field

Table 3: Churn Analysis in Various Business Field

Industry	Application	Publishing Information	Industry	Application	Publishing Information
Game	Churn Prevention	Runge, Julian, et al. (2014) [45]	Game	Game Design	Drachen, Anders, Magy Seif El-Nasr, and Alessandro Canossa, eds. (2013) [72]
Game	Churn Prevention	Periáñez, África, et al. (2016) [14]	Game	Churn Prevention	Kim, Seungwook, et al. (2017) [69]
Game	Churn Prevention, Customer Clustering	Bindewald, Jason M., Gilbert L. Peterson, and Michael E. Miller. (2016) [70]	Game	Churn Prevention	Ben Lewis-Evans. (2012) [71]
Game	Churn Prevention	Lee, Eunjo, et al. (2018) [10]	Game	Churn Prevention	Bertens, Paul, Anna Guitart, and África Periáñez. (2017) [73]
Game	Churn Prevention	Tamassia, Marco, et al. (2016) [75]	Game	Churn Prevention	Hadiji, Fabian, et al. (2014) [42]
Game	Churn Prevention	Bastiaan van der Palen. (2017) [76]	Game	Churn Prevention, Customer Profitability	Sifa, Rafet, et al. (2015) [52]
Game	Churn Prevention, Customer Profitability	Lee, Eunjo, et al. (2018) [17]	Game	Game Design	Fields, Tim, and Brandon Cotton. (2011) [77]
Game	Churn Prevention	Sifa, Rafet, Christian Bauckhage, and Anders Drachen. (2014) [56]	Game	Churn Prevention	Yang, Wanshan, et al. (2019) [43]
Game	Game Design	Fields, Tim. (2014) [3]	Game	Churn, Customer Clustering	Kawale, Jaya, Aditya Pal, and Jaideep Srivastava. (2009) [78]
Game	Churn Prevention	Jeff Grubb. (2014) [39]	Game	Churn Prevention	Kristensen, Jeppe Theiss, and Paolo Burelli. (2019) [79]
Game	Churn Prevention	Nozhnin, Dmitry. (2012) [74]	Game	Churn Prevention	Guitart, Anna, Pei Pei Chen, and África Periáñez. (2019) [80]
Game	Churn Prevention	Viljanen, Markus, et al. (2017) [81]	Game	Churn Prevention	Lee, eunjo. (2019) [82]
Game	Churn Prevention	Bauckhage, Christian, et al. (2012) [83]	Game	Churn Prevention	Shores, Kenneth B., et al. (2014) [84]
Game	Churn Prevention	Debeauvais, Thomas, et al. (2011) [85]	Game	Churn, Customer Profitability	Milošević, Miloš, Nenad Živić, and Igor Andjelković. (2017) [44]
Game	Churn Prevention	Viljanen, Markus, et al. (2016) [40]	Game, Marketing	Churn Prevention	Castro, Emiliano G., and Marcos SG Tsuzuki. (2015) [86]
Game, Psychology	Churn Prevention, Customer Clustering	Borbora, Zoheb, et al. (2011) [47]	Game, Psychology	Churn Prevention, Customer Clustering	Yee, Nick. (2016) [48]
Game, Patent	Churn Prevention	Wolters, Hans, Jim Baer, and Girish Keswani. (2014) [87]	Game, Gamble	Churn Prevention	Coussement, Kristof, and Koen W. De Bock. (2013) [88]

Table 3 Churn Analysis in Various Business Field (*continued*)

Industry	Application	Publishing Information	Industry	Application	Publishing Information
Finance	Churn Prevention	Glady, Nicolas, Bart Baesens, and Christophe Croux. (2009) [38]	Finance	Churn Prevention	Zopounidis, Constantin, Maria Mavri, and George Ioannou. (2008) [37]
Finance	Churn Prevention	Chiang, Ding-An, et al. (2003) [35]	Finance	Churn Prevention	Larivière, Bart, and Dirk Van den Poel. (2004) [36]
Finance	Churn Prevention	Xie, Yaya, et al. (2009) [91]	Finance	Churn Prevention	Anil Kumar, Dudyala, and Vadlamani Ravi. (2008) [92]
Finance	Churn Prevention	Nie, Guangli, et al. (2011) [93]	Finance	Churn Prevention	Ali, Özden Gür, and Umut Arıttürk. (2014) [94]
Finance	Customer Behavior	Evermann, Joerg, Jana-Rebecca Rehse, and Peter Fettke. (2017) [18]	Finance	Customer Profitability	Athanassopoulos (2000) [138]
Finance	Churn Prevention	Lim, Tong-Ming, and Angela Siew Hoong Lee. (2017) [152]	Finance, Telecom., Newspapers	Churn Prevention	Burez, Jonathan, and Dirk Van den Poel. (2009) [95]
Finance	Churn Prevention	De Bock, Koen W., and Dirk Van den Poel. (2011) [103]	Finance, Telecom., Commerce	Churn Prevention, Customer Profitability	Verbraken, Thomas, Wouter Verbeke, and Bart Baesens. (2012) [55]
Finance	Churn Prevention, Customer Profitability	Chu, Tsai, and Ho (2007) [137]	Insurance	Churn Prevention	Morik, Katharina, and Hanna Köpcke. (2004) [98]
Insurance	Churn Prevention	Hur, Yeon, and Sehun Lim. (2005) [99]	Insurance	Churn Prevention	Risselada, Hans, Peter C. Verhoef, and Tammo HA Bijmolt. (2010) [100]
Insurance	Churn Prevention	Zhang, Rong, et al. (2017) [11]			
Marketing	Churn Prevention	Xie, Hanting, et al. (2015) [127]	Marketing	Churn Prevention	Fader, Peter S., Bruce GS Hardie, and Ka Lok Lee. (2005) [49]
Management	Churn Prevention, Customer Profitability	Ngai, Eric WT, Li Xiu, and Dorothy CK Chau. (2009) [30]	Management	Churn Prevention	Lejeune, Miguel APM. (2001) [31]
Management	Churn Prevention, Customer Behavior Analysis	Butgereit, Laurie. (2020) [89]	Management, Patent	Churn Prevention, Customer Profitability	Wright, Christine. (2003) [90]
Retails	Churn Prevention	Clemente, M., V. Giner-Bosch, and S. San Matías. (2010) [97]	Retails	Churn Prevention	Buckinx, Wouter, and Dirk Van den Poel. (2005) [96]
Newspapers	Churn Prevention, Customer Profitability	Neslin, Scott A., et al. (2006) [54]	Newspapers	Churn Prevention	Coussement, Kristof, and Dirk Van den Poel. (2008) [67]
Newspapers	Churn Prevention	Coussement, Kristof, Dries F. Benoit, and Dirk Van den Poel. (2010) [101]	Music Streaming Service	Churn Prevention	Chen, Yian, et al. (2018) [15]

Table 3 Churn Analysis in Various Business Field (*continued*)

Industry	Application	Publishing Information	Industry	Application	Publishing Information
Music Streaming Service	Churn Prevention	Nimmagadda, Sravya, Akshay Subramaniam, and Man Long Wong. (2017) [58]	Music Streaming Service	Churn Prevention	Chen, Min. (2019) [104]
Internet Service	Churn Prevention	Dechant, Andrea, Martin Spann, and Jan U. Becker. (2019) [46]	Internet Service	Churn Prevention	Ngonmang, Blaise, Emmanuel Viennet, and Maurice Tchuenta. (2012) [105]
Internet Service	Churn Prevention	Dror, Gideon, et al. (2012) [57]	Internet Service	Churn Prevention	Madden, Savage, and Coble-Neal (1999) [133]
Internet Commerce	Churn Prevention	Yu, Xiaobing, et al. (2011) [106]	Psychology	Churn Prevention, Customer Profitability	Tamaddoni, Ali, Stanislav Stakhovych, and Michael Ewing. (2016) [102]
Energy	Churn Prevention	Moeyersoms, Julie, and David Martens. (2015) [51]	Pay TV	Churn Prevention	Burez and van den Poel (2007) [131]
Pay TV	Churn Prevention	Burez and van den Poel (2008) [132]	Human Resources	Churn Prevention, Customer Profitability	Saradhi, V. Vijaya, and Girish Keshav Palshikar. (2011) [50]
Telecom.	Churn Prevention	Dalvi, Preeti K., et al. (2016) [107]	Telecom.	Churn Prevention	Hung, Shin-Yuan, David C. Yen, and Hsiu-Yu Wang. (2006) [32]
Telecom.	Churn Prevention	Ge, Yizhe, et al. (2017) [108]	Telecom.	Churn Prevention	Ahn, Jae-Hyeon, Sang-Pil Han, and Yung-Seop Lee. (2006) [33]
Telecom.	Churn Prevention, Customer Profitability	Mozier, Michael C., et al. (2000) [16]	Telecom.	Churn Prevention, Customer Profitability	Wei, Chih-Ping, and I-Tang Chiu. (2002) [41]
Telecom.	Churn Prevention, Customer Profitability	Dahiya, Kiran, and Surbhi Bhatia. (2015) [65]	Telecom.	Churn Prevention, Customer Profitability	Vafeiadis, Thanasis, et al. (2015) [8]
Telecom.	Churn Prevention, Customer Profitability	Verbraken, Thomas, Wouter Verbeke, and Bart Baesens. (2014) [4]	Telecom.	Churn Prevention, Customer Profitability	Baumann, Annika, et al. (2015) [109]
Telecom.	Churn Prevention, Customer Profitability	Dasgupta, Koustuv, et al. (2008) [110]	Telecom.	Churn Prevention	Au, Wai-Ho, Keith CC Chan, and Xin Yao. (2003) [34]
Telecom.	Churn Prevention	Tsai, Chih-Fong, and Yu-Hsin Lu. (2009) [111]	Telecom.	Churn Prevention	Tamaddoni Jahromi, Ali, et al. (2010) [20]
Telecom.	Churn Prevention	Hudaib, Amjad, et al. (2015) [112]	Telecom.	Churn Prevention	Óskarsdóttir, María, et al. (2018) [113]
Telecom.	Churn Prevention	Qian, Zhiguang, Wei Jiang, and Kwok-Leung Tsui. (2006) [114]	Telecom.	Churn Prevention	Wangperawong, Artit, et al. (2016) [115]
Telecom.	Churn Prevention	Radosavljevik, Dejan, Peter van der Putten, and Kim Kyllsbech Larsen. (2010) [68]	Telecom.	Churn Prevention	Richter, Yossi, Elad Yom-Tov, and Noam Slonim. (2010) [53]
Telecom.	Churn Prevention	Hadden, John, et al. (2006) [117]	Telecom.	Churn Prevention	Nath, Shyam V., and Ravi S. Behara. (2003) [118]

Table 3 Churn Analysis in Various Business Field (*continued*)

Industry	Application	Publishing Information	Industry	Application	Publishing Information
Telecom.	Churn Prevention	Lemmens, Aurélie, and Christophe Croux. (2006) [119]	Telecom.	Churn Prevention	Huang, Ying, and Tahar Kechadi. (2013) [120]
Telecom.	Churn Prevention	Huang, Bingquan, Mo-hand Tahar Kechadi, and Brian Buckley. (2012) [121]	Telecom.	Churn Prevention	Idris, Adnan, Muhammad Rizwan, and Asifullah Khan. (2012) [122]
Telecom.	Churn Prevention	Dierkes, Torsten, Martin Bichler, and Ramayya Krishnan. (2011) [123]	Telecom.	Churn Prevention	Kim, Kyoungok, Chi-Hyuk Jun, and Jaewook Lee. (2014) [124]
Telecom.	Churn Prevention	Keramati, Abbas, et al. (2014) [125]	Telecom.	Churn Prevention	Verbeke, Wouter, David Martens, and Bart Baesens. (2014) [126]
Telecom.	Churn Prevention, Customer Profitability	Bahnsen, Alejandro Correa, Djamila Aouada, and Björn Ottersten. (2015) [66]	Telecom.	Churn Prevention	Gerpott et al. (2001) [134]
Telecom.	Churn Prevention	Seo et al. (2008) [135]	Telecom.	Churn Prevention	Pendharkar (2009) [136]
Telecom.	Churn Prevention	Kim and Yoon (2004) [139]	Telecom.	Churn Prevention	Amin, Adnan, et al. (2016) [144]
Telecom.	Churn Prevention	Gui, Chun. (2017) [145]	Telecom.	Churn Prevention	Xia, Guo-en, and Weidong Jin. (2008) [39]
Telecom.	Churn Prevention	Kirui, Clement, et al. (2013) [116]			

B Churn Prediction Models

B.1 Churn Prediction Models

Table 4: Churn Prediction Models

Model Summary	Prediction Algorithms	Best	Publishing Information	Model Summary	Prediction Algorithms	Best	Publishing Information
Traditional Machine Learning	AdaBoost, DT, NB	AdaBoost	Kawale, Jaya, Aditya Pal, and Jaideep Srivastava. (2009) [78]	Traditional Machine Learning	Apriori Algorithm, SVM, NB	SVM	Morik, Katharina, and Hanna Köpcke. (2004) [98]
Traditional Machine Learning	Association Rule	Association Rule	Chiang, Ding-An, et al. (2003) [35]	Traditional Machine Learning	Bagging, NN, Logit, ADT, BN, RF, SVM	NN	Verbraken, Thomas, Wouter Verbeke, and Bart Baesens. (2012) [55]
Traditional Machine Learning	Bagging, RF, RSM, CART, Rotation Forest	Rotation Forest	De Bock, Koen W., and Dirk Van den Poel. (2011) [103]	Traditional Machine Learning	Boosting, Bagging, Logit	Boosting	Lemmens, Aurélie, and Christophe Croux. (2006) [119]
Traditional Machine Learning	Boosting, Logit, DT, NN, RF	Boosting	Clemente, M., V. Giner-Bosch, and S. San Matías. (2010) [97]	Traditional Machine Learning	Boosting, Logit, SVM	Boosting	Tamaddoni, Ali, Stanislav Stakhovych, and Michael Ewing. (2016) [102]
Traditional Machine Learning	CART	CART	Tamaddoni Jahromi, Ali, et al. (2010) [20]	Traditional Machine Learning	DCES, LR, ANN, SVM, NB	DCES	Baumann, Annika, et al. (2015) [109]
Traditional Machine Learning	DT	DT	Wei, Chih-Ping, and I-Tang Chiu. (2002) [41]	Traditional Machine Learning	DT	DT	Richter, Yossi, Elad Yom-Tov, and Noam Slonim. (2010) [53]
Traditional Machine Learning	DT, GAM, RF	RF + GAM ensemble	Coussement, Kristof, and Koen W. De Bock. (2013) [88]	Traditional Machine Learning	DT, Logit, KNN, SVM, K-means Clustering	K-means Clustering	Huang, Ying, and Tahar Kechadi. (2013) [120]
Traditional Machine Learning	DT, Logit, SVM	SVM, DT	Xie, Hanting, et al. (2015) [127]	Traditional Machine Learning	DT, Logit		Dalvi, Preeti K., et al. (2016) [107]
Traditional Machine Learning	DT, LR	DT	Dahiya, Kiran, and Surbhi Bhatia. (2015) [65]	Traditional Machine Learning	DT, LR, NN, NB	DT	Hadiji, Fabian, et al. (2014) [42]
*Abbreviations were defined in Appendix C.							

Table 4 Churn Prediction Models (*continued*)

Model Summary	Prediction Algorithms	Best	Publishing Information	Model Summary	Prediction Algorithms	Best	Publishing Information
Traditional Machine Learning	DT, LR, RF	DT	Bahnsen, Alejandro Correa, Djamil Aouada, and Björn Ottersten. (2015) [66]	Traditional Machine Learning	DT, NB, BN	DT	Kirui, Clement, et al. (2013) [116]
Traditional Machine Learning	DT, NN, LR, CART	DT	Hadden, John, et al. (2006) [117]	Traditional Machine Learning	GAM, LR	GAM	Coussement, Kristof, Dries F. Benoit, and Dirk Van den Poel. (2010) [101]
Traditional Machine Learning	K-means, NMF, PCA, SIVM, AA	AA	Bindewald, Jason M., Gilbert L. Peterson, and Michael E. Miller. (2016) [70]	Traditional Machine Learning	KNN	KNN	Castro, Emiliano G., and Marcos SG Tsuzuki. (2015) [86]
Traditional Machine Learning	KNN	KNN	Borbora, Zohab, et al. (2011) [47]	Traditional Machine Learning	KNN, DT, RF, Logit	KNN	Bastiaan van der Palen. (2017) [76]
Traditional Machine Learning	Logit	Logit	Shores, Kenneth B., et al. (2014) [84]	Traditional Machine Learning	NN, SVM, DT	SVM + Adaboost	Vafeiadis, Thanasis, et al. (2015) [8]
Traditional Machine Learning	Logit, LR, NB, MLP, SVM	Logit, SVM	Huang, Bingquan, Mohand Tahar Kechadi, and Brian Buckley. (2012) [121]	Traditional Machine Learning	Logit, MLN	MLN	Dierkes, Torsten, Martin Bichler, and Ramayya Krishnan. (2011) [123]
Traditional Machine Learning	Logit, RF, PCA, XGB	XGB	Ge, Yizhe, et al. (2017) [108]	Traditional Machine Learning	NB, DT, KNN, GB, Adaboost, Logit, RF, HMM	HMM	Tamassia, Marco, et al. (2016) [75]
Traditional Machine Learning	Logit, DT	Logit, DT	Ali, Özden Gür, and Umut Arıttürk. (2014) [94]	Traditional Machine Learning	Logit, DT	Logit	Risselada, Hans, Peter C. Verhoef, and Tammo HA Bijmolt. (2010) [100]
Traditional Machine Learning	Logit, SVM, C45	SVM	Moeyersoms, Julie, and David Martens. (2015) [51]	Traditional Machine Learning	LR, DT	LR	Nie, Guangli, et al. (2011) [93]

*Abbreviations were defined in Appendix C.

Table 4 Churn Prediction Models (*continued*)

Model Summary	Prediction Algorithms	Best	Publishing Information	Model Summary	Prediction Algorithms	Best	Publishing Information
Traditional Machine Learning	LR, DT	DT	Milošević, Miloš, Nenad Živić, and Igor Andjelković. (2017) [44]	Traditional Machine Learning	LR, DT	LR	Radosavljević, Dejan, Peter van der Putten, and Kim Kylesbech Larsen. (2010) [68]
Traditional Machine Learning	LR, DT, NN, AdaBoost, CSDT	AdaBoost, CSDT	Glady, Nicolas, Bart Baesens, and Christophe Croux. (2009) [38]	Traditional Machine Learning	LR, DT, NN, BN	DT	Neslin, Scott A., et al. (2006) [54]
Traditional Machine Learning	LR, RF, GB	RF	Burez, Jonathan, and Dirk Van den Poel. (2009) [95]	Traditional Machine Learning	LR, SVM, DT, RF	LR	Yang, Wanshan, et al. (2019) [43]
Traditional Machine Learning	NB, Logit, TAN, Max-Min Hill Climbing	Logit	Verbraken, Thomas, Wouter Verbeke, and Bart Baesens. (2014) [4]	Traditional Machine Learning	NB, DT, NN	NN	Nozhnin, Dmitry. (2012) [74]
Traditional Machine Learning	NB, Logit, SVM, DT, RF, KNN	RF	Dror, Gideon, et al. (2012) [57]	Traditional Machine Learning	NB	NB	Nath, Shyam V., and Ravi S. Behara. (2003) [118]
Traditional Machine Learning	DT, Logit	Logit	Lim, Tong-Ming, and Angela Siew Hoong Lee. (2017) [152]	Traditional Machine Learning	NN, DMEL	DMEL	Au, Wai-Ho, Keith CC Chan, and Xin Yao. (2003) [34]
Traditional Machine Learning	NN, DT, SVM, ESVM	ESVM	Yu, Xiaobing, et al. (2011) [106]	Traditional Machine Learning	NN, K-means Clustering + NN, SOM + NN	SOM + NN	Hudaib, Amjad, et al. (2015) [112]
Traditional Machine Learning	NN, Logit, DT, SVM	NN	Runge, Julian, et al. (2014) [45]	Traditional Machine Learning	NN, Logit	NN	Mozier, Michael C., et al. (2000) [16]
Traditional Machine Learning	NN, LR, RF	RF	Buckinx, Wouter, and Dirk Van den Poel. (2005) [96]	Traditional Machine Learning	NN, SOM	NN	Tsai, Chih-Fong, and Yu-Hsin Lu. (2009) [111]
Traditional Machine Learning	SVM	SVM	Hur, Yeon, and Sehun Lim. (2005) [99]	Traditional Machine Learning	SVM, MLP, DT, NN	NN	Keramati, Abbas, et al. (2014) [125]
*Abbreviations were defined in Appendix C.							

Table 4 Churn Prediction Models (*continued*)

Model Summary	Prediction Algorithms	Best	Publishing Information	Model Summary	Prediction Algorithms	Best	Publishing Information
Traditional Machine Learning	SVM, RF, NB	SVM	Saradhi, V. Vijaya, and Girish Keshav Palshikar. (2011) [50]	Traditional Machine Learning	SVM, RF, Logit	RF	Coussement, Kristof, and Dirk Van den Poel. (2008) [67]
Traditional Machine Learning	RF, DT, SVM	RF	Sifa, Rafet, et al. (2015) [52]	Traditional Machine Learning	RF, KNN	RF	Idris, Adnan, Muhammad Rizwan, and Asifullah Khan. (2012) [122]
Traditional Machine Learning	RF, NN, DT, SVM	RF	Xie, Yaya, et al. (2009) [91]	Traditional Machine Learning	RF, SVM, LR, DT, MLP	RF	Anil Kumar, Dudyala, and Vadlaman Ravi. (2008) [92]
Traditional Machine Learning	XGB, NN, Logit	XGB	Nimmagadda, Sravya, Akshay Subramaniam, and Man Long Wong. (2017) [58]	Traditional Machine Learning	K-means Clustering, DT, NN	NN	Hung, Shin-Yuan, David C. Yen, and Hsiu-Yu Wang. (2006) [32]
Traditional Machine Learning	RF, XGB, GBM	RF	Lee, Eunjo, et al. (2018) [17]	Traditional Machine Learning	RF, XGB, LR	RF	Lee, Eunjo. (2019) [82]
Traditional Machine Learning	SVM	SVM	Ngonmang, Blaise, Emmanuel Viennet, and Maurice Tchunte. (2012) [105]	Deep Learning	DNN + LSTM	DNN + LSTM	Guitart, Anna, Pei Pei Chen, and África Periañez. (2019) [80]
Deep Learning	Logit, GB, RF, LSTM	GB	Kim, Seung-wook, et al. (2017) [69]	Deep Learning	DSM, CNN, LSTM, SGD, GB, RF	DSM	Zhang, Rong, et al. (2017) [11]
Deep Learning	CNN	CNN	Mishra, Abinash, and U. Srinivasulu Reddy. (2017) [129]	Deep Learning	CNN	CNN	Umayaparvathi, V., and K. Iyakutti. (2017) [130]
Deep Learning	LSTM	LSTM	Evermann, Joerg, Jana-Rebecca Rehse, and Peter Fettke. (2017) [18]	Deep Learning	RF, NN, LSTM, Hidden State LSTM	Hidden State LSTM	Kristensen, Jeppe Theiss, and Paolo Burelli. (2019) [79]
Deep Learning	CNN + Autoencoders	CNN + Autoencoders	Wangperawong, Artit, et al. (2016) [115]	Graph Theory	Network Analysis (Centrality), MLP	Network Analysis	Kim, Kyoun-gok, Chi-Hyuk Jun, and Jae-wook Lee. (2014) [124]

*Abbreviations were defined in Appendix C.

Table 4 Churn Prediction Models (*continued*)

Model Summary	Prediction Algorithms	Best	Publishing Information	Model Summary	Prediction Algorithms	Best	Publishing Information
Graph Theory	Graph (CDRN, NLB, SPA RC, WVRN), DT, Bagging, RF, BN, Logit	Graph	Verbeke, Wouter, David Martens, and Bart Baesens. (2014) [126]	Graph Theory	Graph Energy distribution	Graph Energy distribution	Dasgupta, Koustuv, et al. (2008) [110]
Interview	User Behavior Interview	-	Ben Lewis-Evans. (2012) [71]				
Statistics	Cox Regression (Survival Analysis)	Survival Analysis	Bertens, Paul, Anna Guitart, and África Periañez. (2017) [73]	Statistics	Cox Regression (Survival Analysis)	Survival Analysis	Viljanen, Markus, et al. (2017) [81]
Statistics	Survival Analysis	Survival Analysis	Bauckhage, Christian, et al. (2012) [83]	Statistics	DTW, Nearest Neighbor, SF	SF + ED, SF + DTW	Óskarsdóttir, María, et al. (2018) [113]
Statistics	Survival Analysis	Survival Analysis	Periañez, África, et al. (2016) [14]	Statistics	Survival Analysis	Survival Analysis	Lee, Eunjo, et al. (2018) [10]
Statistics	Survival Analysis	Survival Analysis	Sifa, Rafet, Christian Bauckhage, and Anders Drachen. (2014) [56]	Statistics	Survival Analysis	Survival Analysis	Zopounidis, Constantin, Maria Mavri, and George Ioannou. (2008) [37]
Statistics	Survival Analysis	Survival Analysis	Larivière, Bart, and Dirk Van den Poel. (2004) [36]	Statistics	Survival Analysis	Survival Analysis	Viljanen, Markus, et al. (2016) [40]
Statistics	Survival Analysis	Survival Analysis	Dechant, Andrea, Martin Spann, and Jan U. Becker. (2019) [46]	Statistics	Time Series Features with Survival Analysis	Time Series Features with Survival Analysis	Qian, Zhiguang, Wei Jiang, and Kwok-Leung Tsui. (2006) [114]
Statistics	Status Diagram with Hypothesis Test	Status Diagram with Hypothesis Test	Ahn, Jae-Hyeon, Sang-Pil Han, and Yung-Seop Lee. (2006) [33]				

*Abbreviations were defined in Appendix C.

C List of Abbreviations

Table 5: List of Abbreviations

Abbreviation	Definition	Abbreviation	Definition
PCA	Principal Component Analysis	NMF	Non-negative Matrix Factorization
NB	Naive Bayes	RF	Random Forest
SF	Similarity Forest	DT	Decision Tree
GAM	Generalized Additive Model	ADT	Alternating Decision Tree
CSDT	Cost Sensitive Decision Tree	CART	Classification and Regression Tree
TAN	Tree Augmented Naive bayes	SVM	Support Vector Machine
ESVM	Evolutionary Support Vector Machine	DMEL	Data Mining by Evolutionary Learning
SOM	Self-Organizing Maps	RSM	Random Subspace Method
Logit	Logistic Regression	LR	Linear Regression
KNN	K-Nearest Neighbors	DCES	Decision Centric Ensemble Selection
GB	Gradient Boosting	XGB	eXtreme Gradient Boosting
HMM	Hidden Markov Model	MLP	Multilayer Perceptron
BN	Bayesian Network	MLN	Markov Logic Network
SGD	Stochastic Gradient Descent	NN	Neural Network (shallow)
ANN	Artificial Neural Network (shallow)	DNN	Deep Neural Network
CNN	Convolutional Neural Network	LSTM	Long Short-Term Memory model
DSM	Deep Shallow mixture Model	DTW	Dynamic Time Wrapping
AA	Archetypal Analysis	ARIMA	Autoregressive Integrated Moving Average
NLB	Network-only Link Based classifier	CDRN	Class-Distribution Relational Neighbor classifier
WVRN	Weighted-Vote Relational Neighbor	SPA-RC	SPatial Association Rule-based Classifier

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., Àngela 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] Perriñez, 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] Larivière, 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] Glady, 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] Hadiji, 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] Saradhi, 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] Dahiya, 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 Periañez. "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 Periañez. "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]I804:11009-000000084744. (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 Arıtürk. "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, et al. “Social ties and their relevance to churn in mobile telecom networks.” *Proceedings of the 11th international conference on Extending database technology: Advances in database technology*. 2008.
- [111] Tsai, Chih-Fong, and Yu-Hsin Lu. “Customer churn prediction by hybrid neural networks.” *Expert Systems with Applications* 36.10 (2009): 12547-12553.
- [112] Hudaib, Amjad, et al. “Hybrid data mining models for predicting customer churn.” *International Journal of Communications, Network and System Sciences* 8.05 (2015): 91.
- [113] Óskarsdóttir, María, et al. “Time series for early churn detection: Using similarity based classification for dynamic networks.” *Expert Systems with Applications* 106 (2018): 55-65.
- [114] Qian, Zhiguang, Wei Jiang, and Kwok-Leung Tsui. “Churn detection via customer profile modelling.” *International Journal of Production Research* 44.14 (2006): 2913-2933.
- [115] Dingli, Alexiei, Vincent Marmara, and Nicole Sant Fournier. “Comparison of deep learning algorithms to predict customer churn within a local retail industry.” *International journal of machine learning and computing* 7.5 (2017): 128-132.
- [116] Kirui, Clement, et al. “Predicting customer churn in mobile telephony industry using probabilistic classifiers in data mining.” *International Journal of Computer Science Issues (IJCSI)* 10.2 Part 1 (2013): 165.
- [117] Hadden, John, et al. “Churn prediction: Does technology matter.” *International Journal of Intelligent Technology* 1.2 (2006): 104-110.
- [118] 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.
- [119] Lemmens, Aurélie, and Christophe Croux. “Bagging and boosting classification trees to predict churn.” *Journal of Marketing Research* 43.2 (2006): 276-286.

- [120] Huang, Ying, and Tahar Kechadi. "An effective hybrid learning system for telecommunication churn prediction." *Expert Systems with Applications* 40.14 (2013): 5635-5647.
- [121] Huang, Bingquan, Mohand Tahar Kechadi, and Brian Buckley. "Customer churn prediction in telecommunications." *Expert Systems with Applications* 39.1 (2012): 1414-1425.
- [122] 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.
- [123] 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.
- [124] 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, Hanting, 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.
- [140] Miglautsch, John R. "Thoughts on RFM scoring." *Journal of Database Marketing & Customer Strategy Management* 8.1 (2000): 67-72.
- [141] O'Brien, Louise, and Charles Jones. "Do rewards really create loyalty?." *Long range planning* 28.4 (1995): 130-130.
- [142] Schweidel, David A., and George Knox. "Incorporating direct marketing activity into latent attrition models." *Marketing Science* 32.3 (2013): 471-487.

- [143] Zhu, Lingxue, and Nikolay Laptev. "Deep and confident prediction for time series at uber." *2017 IEEE International Conference on Data Mining Workshops (ICDMW)*. IEEE, 2017.
- [144] Amin, Adnan, et al. "Comparing oversampling techniques to handle the class imbalance problem: A customer churn prediction case study." *IEEE Access* 4 (2016): 7940-7957.
- [145] Gui, Chun. "Analysis of imbalanced data set problem: The case of churn prediction for telecommunication." *Artificial Intelligence Research* 6.2 (2017): 93.
- [146] Reinartz, Werner J., and Vijay Kumar. "On the profitability of long-life customers in a noncontractual setting: An empirical investigation and implications for marketing." *Journal of marketing* 64.4 (2000): 17-35.
- [147] Fader, Peter S., Bruce GS Hardie, and Jen Shang. "Customer-base analysis in a discrete-time noncontractual setting." *Marketing Science* 29.6 (2010): 1086-1108.
- [148] Fader, Peter S., Bruce GS Hardie, and Ka Lok Lee. "'Counting your customers' the easy way: An alternative to the Pareto/NBD model." *Marketing science* 24.2 (2005): 275-284.
- [149] Fader, Peter S., and Bruce GS Hardie. "Probability models for customer-base analysis." *Journal of interactive marketing* 23.1 (2009): 61-69.
- [150] Fader, Peter S., and Bruce GS Hardie. "How to project customer retention." *Journal of Interactive Marketing* 21.1 (2007): 76-90.
- [151] Morrison, Donald G., and David C. Schmittlein. "Generalizing the NBD model for customer purchases: What are the implications and is it worth the effort?." *Journal of Business & Economic Statistics* 6.2 (1988): 145-159.
- [152] Lim, Tong-Ming, and Angela Siew Hoong Lee. "Loyalty Card Membership Challenge: A Study on Membership Churn and their Spending behavior." *Archives of Business Research* 5.6 (2017).