Predicting B2B Customer Churn for Software Maintenance Contracts

Zhuonan ZHANG
Karlsruhe Institute for Technology, Karlsruhe, Germany, zznglende@gmail.com

Ployplearn RAVIVANPONG
Karlsruhe Institute for Technology, Karlsruhe, Germany, ravivanpong@teco.edu

Michael BEIGL
Karlsruhe Institute of Technology, Karlsruhe, Germany, micheal.beigl@kit.edu

Abstract

Customer churn prediction is a well-known application of machine learning and data mining in Customer Relationship Management, which allows a company to predict the probability of its customer churning. In this study, we extended the application of customer churn prediction to the context of software maintenance contract. In addition, we examined the predictive power of economic factors. Random forest, gradient boosting machine, stacking of random forest and gradient boosting machine, XGBoost, and long short-term memory networks were applied. While an ensemble model and XGBoost performed best, macroeconomic variables did not yield statistically significant improvement in any prediction.

Keywords: customer churn prediction, macroeconomic variables, machine learning, software maintenance service

Introduction

In the research area big data analytics for Customer Relationship Management (CRM), one important element is customer retention and churn management. Previous studies have reached consensus about the economic value of customer retention that

1. the costs to win a new customer is much greater than the costs to retain a customer (Dawes & Swailes, 1999); and
2. long-term customers buy more, and if satisfied, may bring new customers (Ganesh, et al., 2000); and
3. losing customers not only leads to loss of revenue, but also increases the cost for attracting new customers (Athanassopoulos, 2000; Colgate & Danaher, 2000).

Consequently, it is important for a company to reduce its customer churn - when customers stop using a service or a product - by accurately and timely identify churning customers, so that it can prioritize, which customers it should retain, and manage them accordingly. This work focuses on developing a predictive model to identify churning customers, specifically in the B2B context of software maintenance contracts.

Software maintenance is a service from a software vendor that typically includes upgrades, supports and other consulting services. Revenues from software maintenance occupy approximately 50% of total revenues for most software vendors. For example, revenues from product support of MicroStrategy amounted to USD296 million (59% of total revenues) in 2018 (MicroStrategy Inc., 2018). Likewise, Software AG earned EUR214.7 million (52% of total revenue) (Software AG, 2019) while Oracle made USD6.8 billion (61% of total revenues) in cloud services and license support (Oracle Corporation, 2019) in 2019 Q2. After purchasing a software license, a company often also signed a software maintenance contract with the software vendor. Since the contracts generate a large share of recurrent revenues, it is therefore in the interest of a software vendor to identify churning customers in advance, in order to minimize revenue loss.

Customer churn predictions in B2B mostly consider only customers' behavior-based variables or service quality indicators, but rarely macroeconomic variables (Ali & Arıtürk, 2014). When making a decision to cancel a service contract, a company usually takes into consideration its costs and benefits of the services as well as its satisfaction with the service providers (Buckinx & Van den Poel, 2005), which can be derived from the its usage and service history. This motivates the incorporation of Recency, Frequency and Monetary (RFM) based variables, relationship duration (Chen, et al., 2015), and variables that reflect satisfaction and service quality in customer churn predictions (Buckinx & Van den Poel, 2005). However, some companies may also consider the expected economic development in their decision making process. For example, they may evaluate the value proposition of their service providers in changing economic conditions and switch to alternative vendors, whose offerings are more suitable to their changing needs (Ali & Arıtürk, 2014). Thus, considering both customers' behaviors and information on economic environment should lead to more accurate churn prediction, as found in case of B2C by (Ali & Arıtürk, 2014; Van den Poel & Lariviere, 2004).

This study presents two contributions. First, we extended the body of knowledge in customer churn prediction to software maintenance contracts, which according to the review of relevant literature, was not yet studied. Consequently, our work set a baseline for an application in this context. Second, we investigated the effect of including macroeconomic variables on the improvement of customer churn prediction, which was rarely studied in B2B context.

Related Works

Previous studies suggested that RFM-based variables and their extension were important in customer churn prediction (Buckinx & Van den Poel, 2005; Chen, et al., 2015; Liu & Shih, 2005; Cheng & Chen, 2009; Yeh, et al., 2009; Chang, et al., 2010; Hosseini, et al., 2010; Liang, 2010; Miguéis, et al., 2012). Among all predictive algorithm, logistic regression is commonly used (Coussement, et al., 2017). Other established and popular techniques include tree-based models, artificial neural networks based models, support vector machines (SVM) (Vafeiadis, et al., 2015) and ensemble techniques. Since churners are normally the minority, the data are often imbalanced. The preferred evaluation metrics are accuracy and recall and the Area Under the Receiver Operator Curve (AUC). The performance of each algorithms seems to depend heavily on feature engineering.

In B2B context, Chen, Hu and Hsieh (2015) studied customer churn prediction in logistics industry, based on an extended RFM model. They found that decision tree yielded the highest accuracy (93.1%), comparing to multi-layer perceptron, support vector machine (SVM) and logistic regression. Relationship length, recency and monetary variables were found to be important indicators. In addition to the RMF-model, Jahromi, Stakhovych and Ewing (2014) used variables related to efficient retention campaigns. Boosting technique performed better than decision tree and logistic regression, with the AUC of 0.92, in their study.

More algorithms were explored in the B2C context. The case study in automobile and credit card industry from Sundarkumar and Ravi (2015) showed that, after using a hybrid undersampling, decision tree and SVM (AUC = 0.81) performed best (AUC = 0.79). Coussement, Benoit and Van den Poel (2010) employed a Generalized Additive Model (GAM), a flexible extension of logistic regression. They found that GAM could improve the AUC of the baseline logistic regression to 0.85 because it can fit non-linear relationship among the data better. In the latter case study in online gambling, Coussement and De Bock

(2013) confirmed that ensemble learners yielded better and more robust predictions than single learners. An ensemble algorithm based on GAM, combining with bagging and random subspace method (GAMens) were very competitive with random forest (RF) in their study: the top-decile lift (TDL) on out of period data set is 3.47 for RF and 3.40 for GAMens, compared to 2.96 for decision tree and 2.75 of GAM. According to Kasiran et al. (2014) and Simon-Constantinescu et al. (2018), generated good prediction with 97% accuracy for telecommunication industry and 85% recall in retail industry respectively.

Data pre-processing and feature engineering play an important role in customer churn modelling. Churn rate were very low in some cases, leading to very imbalanced data. Nevertheless, these churning customers could still be a source for a large share of revenues. To overcome the imbalanced data problem, Sundarkumar and Ravi (2015) proposed a hybrid undersampling method that was shown to improve the AUC. Morever, more sophisticated feature engineering could generate better prediction. Jahromi et al. [29] suggested that first clustering customers based RFM variables and then building classification churn model for each cluster resulted in promising performance. Gür Ali and Aritürk (2014) proposed a dynamic framework named multiple training observation per customer from different time periods (MPTD), which improved the churn prediction performance through increasing the training data sample size while decreasing the absolute rarity, and increase the sample density. On the other hand, different data preparation treatment (DPT) could make a commonly used logistic regression to outperform more sophisticated methods without DPT, as shown in the study from Coussement, Lessmann and Verstraeten (2017).

Compare to number of studies in B2C contexts, the previous studies in B2B context are much fewer. Researches focused on mostly in sectors like telecommunication, retail, e-commerce, and logistics, but none on B2B software maintenance service. Typical churn prediction models consider mainly the RMF variables that reflect customer behavior. Macroeconomic data have been so far incorporated only in B2C context by a few studies (Ali & Arıtürk, 2014; Van den Poel & Lariviere, 2004). Both used survival analysis and indicated that the GDP growth was valuable for churn prediction. However, other studies did not consider the development of macroeconomic variables over time. There still exists a gap in the case of software maintenance service under B2B context, and an investigation on the impact of economic variables in the models, which is examined in this study.

Methods

Data

The sampled data set we used in this case study consisted anonymized CRM (contract history, master data, tickets, satisfaction) data, which are RFM-based, and macroeconomic data. The data contain information of over 25,000 customers, which are small, medium and large companies, with an overall churn rate of less than 5%. So that there was enough historical data, only long term customers, who have been in the contracts for longer than 18 months were considered. Customers' data included information on customer demographics and customers' behavior. Customer demographics are information about customer in business context that usually will not change with time. Most of them are categorical variables e.g. geographical location, market segment, and company size. Customers' Behavior included a company's usage frequency of different software and services, its satisfaction, calculated features based on these information, as well as whether and when a customer cancelled the contract. Usually software maintenance contracts are extended automatically if customers do not cancel them a few months in advance before the end of the contracts. Even though these customers may eventually decide otherwise, the customers, who cancel the contracts, are more likely to stop using the services, and thus are defined as churners in our case. The customer data were aggregated at monthly interval.

Macroeconomic data between 2016 to 2018 were used as a dynamic aspect for churn prediction. Selected variables could be divided into three groups: country-level indicators (GDP growth rate, GDP annual growth rate, consumer spending, consumer confidence, business confidence, gasoline prices), industry-level indicators (GDP manufacturing, service, mining, public administration, GDP utilities, construction, agriculture, and transport, car production statistics, and retail sales), and stock market indicators (stock price, earning per share (EPS)). These indicators were reported at different intervals, ranging from monthly to annually.

Data Preprocessing and Experiment Set Up

The main challenge in this study was the enhancement of RFM-based variables from customer data with economic data as well as feature engineering and selection. So that we did not lose information on customers' behavior, we opted the monthly customer data set as our main data set. Cubic spline interpolation, which is robust and efficient (Ajao, et al., 2012), were used to convert other annually and quarterly data to monthly data.

The combined data set were split into training and holdout data set. To ascertain that the predictive models were robust against future unknown data and to avoid the scenario, in which the model "remembered" the seasonal peaks, the training and the holdout data set had different time distribution. In other words, the holdout data lagged the training data by three months (Figure.1). The training data were used to predict the customer churn in next 3 months, whereas prediction of

holdout data was for next 6 months. Both data set took 18-month historical data as inputs. The training data set were balanced via randomly stratified sampling and were further split into two parts: 85% for training and 15% for validation data sets.

Figure 1: Time series splitting of training and hold-out data set

In order to investigate the effect of economic variables, two groups of models were built: baseline models, which were based only on customer data; and extended models, which incorporated macroeconomic data.

Feature Engineering and Selection

To capture the dynamic aspect and avoiding noise at the same time, variables were aggregated on the predefined period-length level. The value of variables in each period were taken from the latest value. Period mean, maximum, minimum values were calculated. Table 1 summarizes the other major feature engineering based on customers' behavior and economic variables.

Table 1: Feature engineering based on economic variables

Variable	Calculation	Description
V_{mxm}	$\left(V_t - \frac{V_{t-1}+V_{t-2}+\ldots+V_{t-n}}{x}\right) \Big/ \left(\frac{V_{t-1}+V_{t-2}+\ldots+V_{t-x}}{x}\right)$	Development trend comparing to x-month average, where *V* is a variable; *t* is the current month; and $x \in \{6,12,18,24\}$, the predefined period-length.
V_{Δ}	$V_{\max} - V_{\min}$	Maximum – minimum spread in each period
V_{pop}	$\begin{cases} (V_t - V_{t-1})/V_t & \text{if} \quad V_{t-1} \neq 0 \\ V_t - V_{t-1} & \text{if} \quad V_{t-1} = 0 \end{cases}$	Trend between periods
P/E	$\frac{Stock\ Price}{EPS}$	Price per earnings ratio
P/E_{mean}	$\frac{PE_1 + PE_2 + \ldots + PE_n}{N}$	Average price per earning of the sector where PE_n is the *P/E* ratio of a competitor; and *N* is the total number of competitors in the industry.
PE_{ratio}	$\frac{P/E - P/E_{mean}}{P/E_{mean}}$	Performance of the company relative to its competitors

After adding the economic variables, the total number of features increased to 282. Since such high feature dimension would slow down the model training task, increase the computational cost and could make it difficult to find the effective features, feature selection was necessary. Especially irrelevant features, which do not change the target concept learns through machine learning; and redundant features, which do not include anything new to the target concept (John, et al., 1994), should be removed. First, Spearman correlation was employed to filter redundant features among economic variables. Next, K-means algorithm was used to further reduce number of variables. An alternative method would have been principal component analysis (PCA). However, we opted for a clustering algorithm because PCA does not yield disjoint subsets, making it difficult when interpreting feature importance. Considering the time dynamical of features, some relationship might occur by chance. Thus feature clustering is employed in each period, which may result in different number of clusters. Then Spearman correlation was again employed to combine highly correlated features in each clusters by period together, and features were considered redundant only if it exhibited high correlation in all periods. The overall process in this step could be seen from Figure.2. The categorical features were selected based on χ^2 independence test. Lastly, random forest was employed to select overall effective features, with the cutoff threshold defined by the result of grid search. As a result, 110 features are selected as predictors.

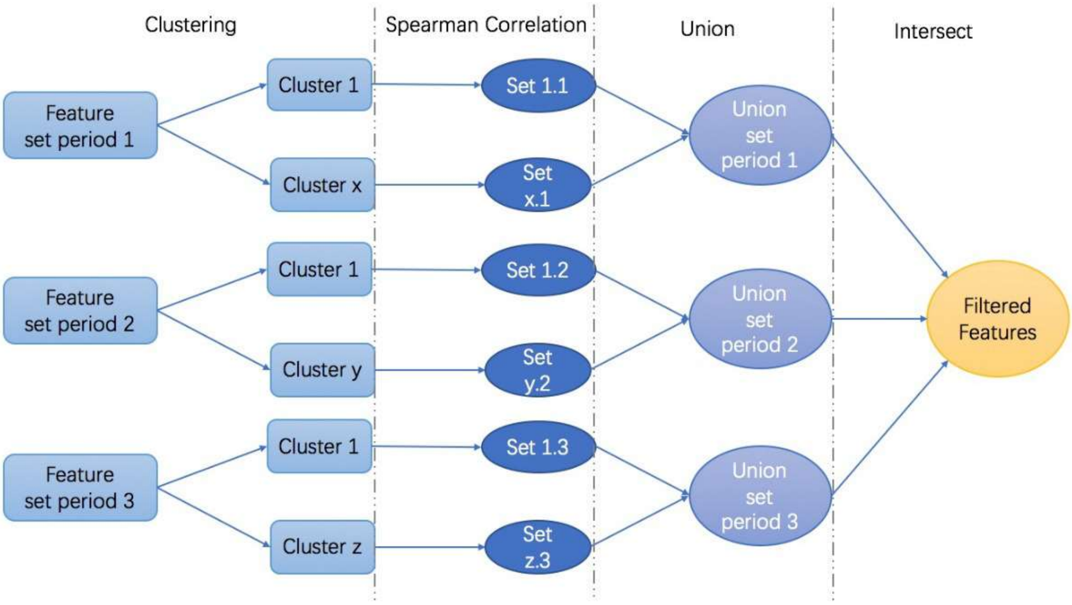

Figure 2: Feature selection based on period

Performance Measure

The AUC is the typical performance metric for churn prediction. In contrast to some other evaluation measurements e.g. accuracy), it is not influenced by any threshold value, since it takes into account all possible thresholds on the predicted probabilities (Coussement, et al., 2010). Therefore, in this work, AUC was selected as an evaluation metric.

Algorithms and Parameter Settings

Random forest, GAM, GBM, stacked ensemble of random forests and GBMs, and XGBoost were employed. Hyperparameters were optimized based on AUC, using grid search with 5-fold cross validation on the validation data set. Table 2 listed the resulting hyperparameters.

Table 2: Hyperparameter settings, optimized based on AUC on validation set

Classifier	**Hyperparameter**
Random Forest	number of trees = 200; max. depth =20; min. =3; number of bins = 15
GBM	number of trees = 100; max. depth =8; learning rate = 0.05; column sampling rate = 0.2
XGBoost	iteration = 139; max. depth = 20; learning rate = 0.1; column sampling rate = 0.5; $\lambda = 0.2$; $\alpha = 0.1$

For the stacked ensemble, we trained random forests and GBMs with different combination of hyperparameters. The models with top five best AUC were selected as base learners. The superlearner algorithm of the R-package H2O were used, with default settings, to combine the base learners.

Results

In order to examine, which algorithm performs best and whether macroeconomic data could improve prediction, customer churn prediction models were learned using random forest, GAM, GBM, stacking of random forest and GAM, and XGBoost on *(i)* customer data (*baseline models*); and *(ii)* on both customer and macroeconomic data (*extended models*). Table 3 displays the resulting AUC for each model based on holdout data set.

Table 3: AUC of each model on whole holdout data

Algorithm	Baseline	Extended Model
Random forest	**0.6643**	0.6416
GBM	0.6369	0.6366
Ensemble	0.6608	0.6532
GAM	0.6484	**0.6631**
XGBoost	0.6636	**0.6607**

Random forest had highest AUC among baseline models, whereas GAM performed best among the extended models. However, GAM ignored missing values, resulting in no prediction for customer with missing values in the hold-out data set. Therefore, XGBoost was the best model for practical use. All extended models performed similarly. As the matter of fact, most of the baseline models yielded better prediction than extended models.

In order to ascertain if the resulting AUCs were statistically different from one another, we conducted a Wilcoxon test on the hold-out data. The average AUCs of each model were summarized in Table 4. The bold font signified the AUC of an algorithm that was statistically different (p-values < 5%) from the next lowest AUC and were not different among themselves, based on the p-values reported in Table 5 and Table 6. Among the baseline models, random forest, stacked ensemble and XGBoost had similar predictive power. Even though GAM had the highest average AUC, it was not statistically higher than the lowest AUC of the GBM. Analogous result was observed among extended models. In this group, stacked ensemble and XGBoost had comparable performance, which was not statistically different than one another.

Table 4: Average AUCs of baseline and extended models based 10-fold cross-validation

Model	Baseline	Extended Model
Random forest	**0.6647**	0.6421
GBM	0.6370	0.6368
Ensemble	**0.6610**	**0.6535**
GAM	**0.6644**	**0.6611**
XGBoost	0.6696	**0.6715**

Table 5: p-values of Wilcoxon-test among baseline models

		Model 2			
		GBM	**Ensemble**	**XGBoost**	**GAM**
Model 1	**Random Forest**	0.0039	0.2324	0.8457	0.9219
	GBM		0.0039	0.0020	0.2754
	Ensemble			0.6953	0.7695
	XGBoost				0.6250

Table 6: p-values of Wilcoxon-test among extended models

		Model 2			
		GBM	**Ensemble**	**XGBoost**	**GAM**
Model 1	**Random Forest**	0.6250	0.0645	0.0371	0.5566
	GBM		0.0371	0.0273	0.3223
	Ensemble			0.1309	0.7695
	XGBoost				0.6250

Conclusion

In this work, we extended the body of knowledge in customer churn prediction to software maintenance contracts, setting a baseline AUC for an application in this context. Additionally, we investigated the effect of macroeconomic variables on the improvement of customer churn prediction. Our results confirmed the previous studies (Buckinx & Van den Poel, 2005; Chen, et al., 2005; Cheng & Chen, 2009; Liu & Shih, 2005; Chang, et al., 2010; Yeh, et al., 2009; Hosseini, et al., 2010; Liang, 2010; Miguéis, et al., 2012), that variables based on customer behavior plays important role for customer churn prediction. In contrast to the work from Gür Ali and Aritürk (2014), as well as Poel and Lariviere (2004), GDP growth was not a significant predictor. It was ranked 46th among feature importance by XGBoost, the best prediction algorithm among extended models. In fact, including economic variables even worsened the predictive performance in the hold-out data. Based on the average AUC on the hold-out data and the performance stability after adding economic variables, the best algorithms were stacked ensemble of random forests and GBMs (average AUC = 0.6610) and XGBoost (average AUC = 0.6644).

Customers' behavior and its dynamic were significant in predicting customer churn. Ratio of number of software maintenance contracts to number of installed software received the highest weight in most models. The RFM-based variables such as usage history were reported as useful by the XGBoost algorithm, even though a lot of observations were missing. This suggested that cultivating such features should improve prediction performance.

Although the lower AUC on the hold-out data in extended models implied that economic variables did not contribute the any improvement, the worsened performance could be attributed to the fact that they were important to some models during training. Consumer spending, consumer confidence and industrial indicators seemed to improve the AUC in the training data, being selected among top 20 important features. Two possible explanations could be offered for this phenomenon. First, the one stage of feature selection was based on share of missing values. The configuration might not be optimal, so that some effective consumer behavior variables were filtered out while more economic variables were kept. Second, economic indicators did indeed have predictive power during the training period. However, their relationship might have changed in the hold-out period, which lagged the training period by three months, leading to worsen performance. It should be noted that the feature selection process discarded all stock market variables due to lower number of observations. Thus, this study could not provide evidence for the predictive strength of stocks market variables.

Our work demonstrated the application of customer churn prediction for software maintenance contracts. The optimal AUC under our setting was achieved by stacked ensemble and XGBoost, based only customers' data while adding economic indicators worsened the performance. A systematic study on the relationship change between economic indicators and customer churn should be considered. Further prediction improvement could be achieved through a systematic study on more effective feature engineering and feature selection method; although it is also interesting to compare the performance against that of artificial neural networks without feature engineering nor selection.

References

Ajao, I., Ibraheem, A. & Ayoola, F., 2012. Cubic Spline Interpolation: A Robust Method of Disaggregating Annual Data to Quarterly Series. *Journal of Physical Sciences and Environmental Safety*, 2(1), 1-8.

Ali, Ö. G. & Arıtürk, U., 2014. Dynamic churn prediction framework with more effective use of rare event data: The case of private banking. *Expert Systems with Applications*, 41(17), 7889-7903.

Athanassopoulos, A. D., 2000. Customer satisfaction cues to support market segmentation and explain switching behavior. *Journal of business research*, 47(3), 191-207.

Buckinx, W. & Van den Poel, D., 2005. Customer base analysis: partial defection of behaviourally loyal clients in a non-contractual FMCG retail setting. *European Journal of Operational Research*, 164(1), 252-268.

Chang, E.-C., Huang, S.-C. & Wu, H.-H., 2010. Using K-means method and spectral clustering technique in an outfitter's value analysis. *Quality & Quantity*, 44(4), 807-815.

Cheng, C.-H. & Chen, Y.-S., 2009. Classifying the segmentation of customer value via RFM model and RS theory. *Expert systems with applications*, 36(3), 4176-4184.

Chen, K., Hu, Y.-H. & Hsieh, Y.-C., 2015. Predicting customer churn from valuable B2B customers in the logistics industry: a case study. *Information Systems and e-Business Management*, 13(3), 475-494.

Chen, M.-C., Chiu, A.-L. & Chang, H.-H., 2005. Mining changes in customer behavior in retail marketing. *Expert Systems with Applications*, 28(4), 773-781.

Colgate, M. R. & Danaher, P. J., 2000. Implementing a customer relationship strategy: The asymmetric impact of poor versus excellent execution. *Journal of the academy of marketing science*, 28(3), 375-387.

Coussement, K., Benoit, D. F. & Van den Poel, D., 2010. Improved marketing decision making in a customer churn prediction context using generalized additive models. *Expert Systems with Applications*, 37(3), 2132-2143.

Coussement, K. & De Bock, K. W., 2013. Customer churn prediction in the online gambling industry: The beneficial effect of ensemble learning. *Journal of Business Research*, 66(9), 1629-1636.

Coussement, K., Lessmann, S. & Verstraeten, G., 2017. A comparative analysis of data preparation algorithms for customer churn prediction: A case study in the telecommunication industry. *Decision Support Systems*, Band 95, 27-36.

Dawes, J. & Swailes, S., 1999. Retention sans frontieres: issues for financial service retailers. *International*, Band 17, 36-43.

Ganesh, J., Arnold, M. J. & Reynolds, K. E., 2000. Understanding the Customer Base of Service Providers: An Examination of the Differences between Switchers and Stayers. *International Journal of Bank Marketing*, 64(3), 65-87.

Hosseini, S. M. S., Maleki, A. & Gholamian, M. R., 2010. Cluster analysis using data mining approach to develop CRM methodology to assess the customer loyalty. *Expert Systems with Applications*, 37(7), 5259-5264.

Jahromi, A. T., Stakhovych, S. & Ewing, M., 2014. Managing B2B customer churn, retention and profitability. *Industrial Marketing Management*, 43(7), 1258-1268.

John, G. H., Kohavi, R. & Pfleger, K., 1994. Irrelevant features and the subset selection problem. Proceedings of the Eleventh International Conference, New Brunswick, USA, 10–13 July 1994, 121-129.

Kasiran, Z., Ibrahim, Z. & Ribuan, M. S. M., 2014. Customer churn prediction using recurrent neural network with reinforcement learning algorithm in mobile phone users. *International Journal of Intelligent Information Processing*, 5(1), p. 1.

Liang, Y.-H., 2010. Integration of data mining technologies to analyze customer value for the automotive maintenance industry. *Expert systems with Applications*, 37(12), 7489-7496.

Liu, D.-R. & Shih, Y.-Y., 2005. Hybrid approaches to product recommendation based on customer lifetime value and purchase preferences. *Journal of Systems and Software*, 77(2), 181-191.

MicroStrategy Inc., 2018. *MicroStrategy Incoporated 2018 Annual Report.* [Online] [Retrieved August 20, 2019], https://ir.microstrategy.com/financial-information/annual-reports-and-proxies

Miguéis, V. L., Van den Poel, D., Camanho, A. S. & e Cunha, J. F., 2012. Modeling partial customer churn: On the value of first product-category purchase sequences. *Expert systems with applications*, 39(12), 11250-11256.

Oracle Corporation, 2019. *Fourth Quarter Fiscal Year 2019 Press Release.* [Online] [Retrieved August 20, 2019], https://investor.oracle.com/financial-reporting/quarterly-reports/default.aspx

Simion-Constantinescu, A., 2018. *Deep Neural Pipeline for Churn Prediction.* Proceedings of 17th Proceedings of 17th Networking in Education and Research (RoEduNet) Conference, ISBN: 978-1-5386-7135-1, 6-8 September 2018, Cluj-Napoca, Romania, 1-7.

Software AG, 2019. *Software AG Financial Information Q2/2019.* [Online] [Retrieved August 20, 2019], https://investors.softwareag.com/en/result-center

Sundarkumar, G. G. & Ravi, V., 2015. A novel hybrid undersampling method for mining unbalanced datasets in banking and insurance. *Engineering Applications of Artificial Intelligence*, Band 37, 368-377.

Tamaddoni Jahromi, A., Sepehri, M. M., Teimourpour, B. & Choobdar, S., 2010. Modeling customer churn in a non-contractual setting: the case of telecommunications service providers. *Journal of strategic marketing*, 18(7), 587-598.

Vafeiadis, T., Diamantaras, K., Sarigiannidis, G. & Chatzisavvas, K., 2015. A comparison of machine learning techniques for customer churn prediction. *Simulation Modelling Practice and Theory*, June, Band 55, 1-9.

Van den Poel, D. & Lariviere, B., 2004. Customer attrition analysis for financial services using proportional hazard models. *European journal of operational research*, 157(1), 196-217.

Yeh, I.-C., Yang, K.-J. & Ting, T.-M., 2009. Knowledge discovery on RFM model using Bernoulli sequence. *Expert Systems with Applications*, 36(3), 5866-5871.