UKRAINIAN CATHOLIC UNIVERSITY

MASTER THESIS

Customer Lifetime Value For Credit Limit Optimization

Author:
Markiyan KOSTIV

Supervisor:
Oleksandr Romanko

A thesis submitted in fulfillment of the requirements for the degree of Master of Science

in the

Department of Computer Sciences
Faculty of Applied Sciences

Lviv 2019

Declaration of Authorship

I, Markiyan KOSTIV, declare that this thesis titled, "Customer Lifetime Value For Credit Limit Optimization" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

Date:

UKRAINIAN CATHOLIC UNIVERSITY

Abstract

Faculty of Applied Sciences

Master of Science

Customer Lifetime Value For Credit Limit Optimization

by Markiyan KOSTIV

Customer lifetime value is an important metric for banks to optimize a credit limit, improve retention and set competitive pricing. The specifics of credit cards market provide challenges with undetermined usage time and positive correlation between the risk and revenue. To address these challenges, we present a customer lifetime value framework in conjunction with risk-adjusted-return for revolving products and credit limit increase and decrease strategy taking into account CLV metrics.

Acknowledgements

I would like to express gratitude to Queens University and Scotiabank for making this happen, providing computational resources and valuable feedback. A special gratitude I give to Mikhail Nediak and Ivan Mendoza for providing supervision, especially participating in methodology development and providing domain insights.
Also, I would like to thank GRM and Global Lifecycle teams in Scotiabank, Mitacs program, Oleksandr Romanko, Mark Wagner and everyone else who took part in organization of this research internship.

Contents

Chapter 1

Introduction

1.1 Motivation

Customer Lifetime Value (CLV) is becoming a crucial metric for banking to allocate proper resources and invest in potentially profitable customers. Because of the variety of products and changes in customer behavior, it is vital to understand customer needs in perspective. Current credit card limit programs are based on product and rely heavily on short-term risk-based predictions. The business objective of credit limit optimization is to accelerate primary relationships by delivering the most relevant credit offers. The goal of this study is to take limit programs to the customer level and understand customer behavior throughout the whole lifetime. Within this work, we developed a customer lifetime value framework in conjunction with risk-adjusted-return for revolving products and credit limit increase and decrease strategy taking into account CLV metrics.

1.2 Revolving products

Credit is a type of revolving product which can be used for an undetermined amount of time. The amount drawn can fluctuate depending on the customer's cash flow need. Once the debt is repaid, it can be borrowed again. There are two significant challenges with the revolving credits: how to properly optimize the amount of debt to propose and estimate the amount of time a customer is going to stay with the bank. A credit limit defines the maximum amount of credit that a bank is willing to grant to the customer. The essential elements in the definition of the credit limit are the minimization of the risk exposure and maximization of the revenue. The risk is defined as: "the potential that a bank borrower or counterparty will fail to meet its obligations in accordance with agreed terms." [3] . The risk-adjusted return (RAR) framework directly addresses the issue of calculating the return, taking into account risk level. We are not going to treat the RAR calculation, however, it is important to mention since it is one of the inputs to the CLV prediction. The other challenge is to estimate the duration of the relationships with the customer. Since revolving products do not have a determined amount of time it becomes hard to define the usage lifetime. However, it is extremely important because credit issuer company needs to decide on the credit limit plan during the lifetime and resources to retain the customer.

Chapter 2

Related work

Several approaches were studied for years to calculate CLV. Statistical models were limited in data and restricted by computational resources. The progress in machine learning allowed to take a look at this task from another point of view and create more robust predictions.

2.1 RFM models

RFM (Recency, Frequency, Monetary Value)[15] is a method to measure customer value. These models were developed for direct marketing to improve response rates. Prior to these models, companies used demographic data for targeting users; however, it was investigated that monetary value and purchases frequency analysis are better predictors of the future purchases, therefore, better estimate customer value. RFM models create groups based on the three variables:

- Recencey - time since the last purchase was made
- Frequency - frequency of customers' purchases
- Monetary Value - how much do customers' purchases worth

Those groups than can be weighted, so the score is assigned to each customer segment. Despite the fact that RFM models predict the future customer's behavior, they are limited with the next time period. The objective of CLV model is to predict customer's behavior during the whole lifetime. Also, RFM models are limited in variables and ignore that customer's behavior can be affected by the previous marketing activities. Another limitation is that RFM models are predicting the score of the customer and do not provide explicitly the monetary value. Overall, this method provides a good starting point to identify important features, which should be good predictors for customer value.

2.2 Pareto/NBD

One of the most widely used methods to estimate CLV is Pareto model. [8] The model relies on few assumptions about the customer and his or her behavior:

- The customer is considered alive for a certain period of time, after that he or she is defined as permanently inactive
- Customer's purchase frequency follows Poisson distribution, which means that the timing of given purchases is random, but the rate is constant
- The lifetime of customers follows an exponential distribution with dropout rate

The inputs to the model are based on the lifetime duration of the customer, the number of transactions customer has made and the time since the last customer's transaction. This means that the method neglect all other information about the customer such as demographic data or earlier transactions history. Also, the model does not address the monetary value of the predicted transactions.

2.3 Gamma-Gamma Extension

To estimate the monetary value of the transaction was developed an extension to Pareto/NBD model called Gamma-Gamma [22]. To assign a value to each transaction Gamma-Gamma model provides three general assumptions:

- Any customer's transaction value varies randomly around his or her average transaction value
- Average values of transactions range between customers but do not change over time for a particular individual.
- The transaction process does not impact the average transaction distribution

The Gamma-Gamma extension to Pareto/NBD model provides a way to estimate the lifetime of the customer, purchases frequency and overall value. The simplicity of the solution gives high-level interpretability, however, does not take into account any additional information about the customer.

2.4 Econometric models

Econometric models tend to share principles of probabilistic models described above. These models consist of few sub-models for acquisition, retention, and expansion. Having them all combined together we can estimate CLV. Within this work, we are specifically interested in retention component. Customer retention is a probability of a customer being "alive" or repeat buying from a firm[11]. The company has to identify whether the customer is still active and predict the lifetime of the customer. For revolving product, customer informs the bank about the relationships termination so that we can identify the active customers; however, there is no data about the predefined usage time. To estimate customer defection we can build a hazard model. Hazard models can be separated into two groups: AFT (accelerated failure time) and PH (proportional hazard) models. The latter is widely used in the retail industry due to simplicity and ease of estimation.

2.5 RAR model

CLV models consider discounted profits that customer generates over the expected lifetime of relationships with the firm[27][24]. Such approach can be misleading since it ignores the risk. The values of profit and risk are positively correlated in the credit cards market. Assigning a higher limit leads to higher CLV metrics, but at the same time increases the probability of default which might lead to higher losses. Therefore, stakeholders define the risk limits they are able to take and risk-adjusted return (RAR) framework addresses a trade-off between risk and profit. RAR model can incorporate different types of risk such as the probability of default and volatility of revenue streams.
Overall, the treatment of RAR model is out of the scope of this work, however, to address the issue with profit and risk positive correlation we input RAR as a value estimate for each customer for current and future values.

Chapter 3

Proposed model

The goal of this study is to build a customer lifetime value model based on historical data and try to predict long-term revenue behavior. To do that we train a survival model to predict the lifetime expectancy for each customer, divide customers into different segments defined by revenue behavior and model their transition among those groups throughout the whole lifetime. In conjunction with RAR model, we estimate the customer lifetime value.

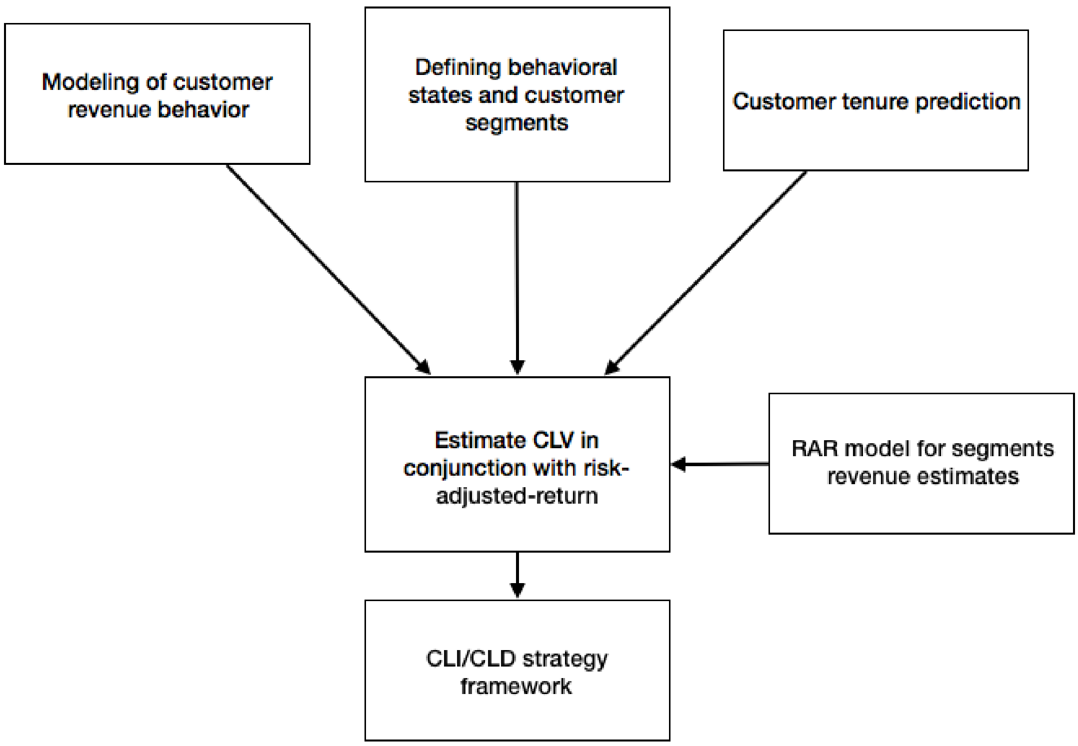

FIGURE 3.1: Flowchart of the proposed model

3.1 Survival model

The purpose of survival analysis [26] is to measure the lifespans of individuals. With this model, we want to predict the time before an event occurs. Depending on the model configuration we can measure the time in different units, such as years, months, days, etc. Even though the problem we trying to solve sounds related to classical regression analysis, survival modeling addresses a crucial issue with censored observations.

3.1.1 Kaplan-Meier survival curves

Kaplan-Meier survival curves [18] is the essential way to estimate the survival function for the population. This non-parametric approach is designed to distinguish between the survival behavior of different groups. The estimator is defined as the difference between the total number of subjects at risk n_i prior time t and the number of events which occurred d_i prior time t divided by the n_i. This method is a great tool for the comparison of survival curves of different groups and provide insights about the general survival distribution [9]. The goal of the survival modeling in terms of CLV framework is to bring it to the customer level and get a precise estimate of subject's lifespan.

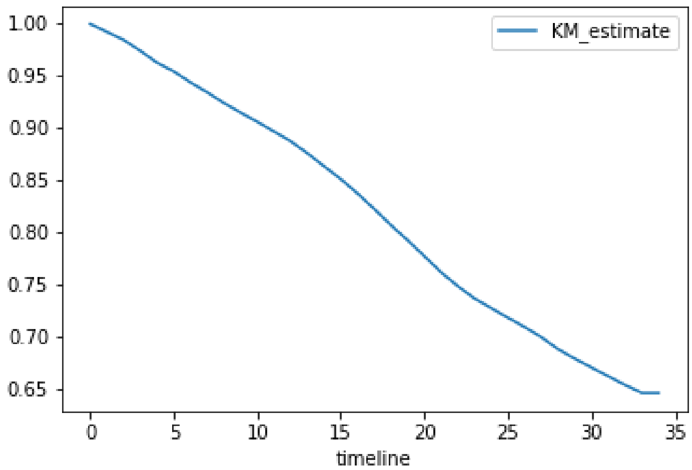

FIGURE 3.2: Survival curve estimated with Kaplan-Meier model

$$\hat{S}(t) = \prod_{t_i t} \frac{n_i - d_i}{n_i} \tag{3.1}$$

3.1.2 Multivariate statistical modeling

There are multiple variables which can influence a lifespan of the customer for the particular product (e.g. purchase frequency, usage of similar products, etc.). The model has to take into account the possible effect of these covariates on the customer's lifespan. The multivariate statistical models are able to predict the risk of survival in respect to the several factors simultaneously.

3.1.3 Cox's Proportional Hazard model

Cox's Proportional Hazard (CPH) model [7] is a method to estimate the impact of several variables upon the time before an event occurred. The hazard function $h(t)$ represents the risk of an event occurring at time t. The major assumption for this method is that baseline hazard h_0 is constant for all the subjects at time t. Cox proportional hazard model contains the time component only in baseline hazard function, which means covariates only can influence the risk prediction, by decreasing or increasing the baseline hazard value.

$$h(t|x) = \overbrace{h_0(t)}^{\text{baseline hazard}} \underbrace{\exp\left(\sum_{i=1}^{n} b_i(x_i - \overline{x_i})\right)}_{\text{partial hazard}} \tag{3.2}$$

The important note about the Cox's Proportional Hazard model is that it allows to build survival curve for each subject in respect to all covariates, however, all survival curves for different subjects have the same basic shape.

3.1.4 Random survival forest

Random survival forest (RSF) [17] is a tree ensemble method designed to capture the nonlinear effects of the covariates on the survival predictions. Unlike the Cox's Proportional Hazard model Random survival forest does not make any assumptions about the baseline hazard function and is more flexible in predicting hazard estimates [20]. The RSF model shares the basic concepts of a classical Random Forest algorithm [13] used for the classification or regression tasks. The RSF model uses decision trees as a base learner. The major difference is the calculation of cumulative hazard function (CHF) for each survival tree and averaging them to get ensemble CHF. [31].

3.1.5 Estimation of RSF survival curves

In contrast to the CPH model, RSF provides the survival probability for the subject over time. To get a survival curve estimates for each subject was used a combination of RSF and CPH models. The survival probability received from the RSF model was used as a single covariate for the CPH model. As a result, we get the assumption of the baseline hazard function and a nonlinear relation between this function and survival probability.

3.1.6 Survival SVM

Support Vector Machines is a classic machine learning algorithm and with health index, [29] modification can be applied to the survival analysis. Due to lack of computation resources, the SVM model was tested only with a linear kernel.

3.1.7 Extrapolation with exponential curve

The survival analysis provides the survival curve only within the observed time period. However, the data which was used for the survival model contains a lot of censored observations and CLV model requires estimation of customer's lifespan. To address this issue we define the maximum possible lifespan M for the entire population based on the historical data and business objective. The survival curve for each customer is than extrapolated to the maximum defined time, so within each time point $t_i \in [0, M)$ survival probability $p_i \in [0, 1]$. The survival curve for each customer use non-linear least squares to find parameters and fit a function:

$$f(x) = a * \exp^{-b*x} + c \tag{3.3}$$

3.2 Evaluation of survival model

To evaluate the survival model were used three different metrics: concordance index [23] Precision and F1-score of customer survival within the observed time period.

3.2.1 Concordance index

The concordance index metrics is one of the essential measurements for the survival models performance. It addresses the issue with censored observations and is independent of the fixed time point of the evaluation. The concordance index is described "as the fraction of all pairs of subjects whose predicted survival times are correctly ordered among all subjects that can actually be ordered" [23]. The concordance index calculation can be expressed as a following mathematical equation [30]. Let $\overline{\eta_j}$ and $\overline{\eta_i}$ be predicted values from the survival model. Then, $\eta_i^1 = \overline{\eta_i}|Y_i = 1$ and $\eta_j^0 = \overline{\eta_j}|Y_j = 0$, where Y is a boolean variable which identifies weather event occured. N_1 and N_2 are the number of event and number of non-events, respectively.

$$C = \frac{1}{N_1 N_2} \sum_{i=1}^{N_1} \sum_{i=1}^{N_2} I(\eta_i^1, \eta_j^0) \tag{3.4}$$

where

$$I(\eta_i^1, \eta_j^0) = \begin{cases} 1 & \text{if } \eta_i^1 > \eta_i^0 \\ 0.5 & \text{if } \eta_i^1 = \eta_i^0 \\ 0 & \text{if } \eta_i^1 < \eta_i^0 \end{cases}$$

The value of $C \in [0, 1]$ corresponds to the ability of the model to distinguish between low and high risk subjects. $C = 1$ interpreted as perfect prediction accuracy and $C = 0.5$ as accuracy of random guess.

3.2.2 Precision and F1-score

To evaluate how robust are predictions of the survival model they could be compared with the historical data. Survival curves for each customer are thresholded with 0.5 value. This means that the maximum lifespan of the customer is defined as a first time point t_i, when the probability of survival p_i is lower than 0.5. Let T be the maximum time point from the training data, than for each customer we define the $\overline{y_i} = t_i \leq T$, which corresponds weather event occurred within the observed time period. Having the predicted $\overline{y}$ value for each customer and ground truth y for customers who have experienced the event, we can calculate the Precision and F1-score metrics. Important to mention that 0.5 threshold is a tunable parameter and depends on the business objective and strategy, so the model performance is measured by concordance index.

TABLE 3.1: Precision, F1-score and concordance index comparison for Cox's Proportional Hazard model, Random Survival Forest and Survival SVM in conjunction with CPH

Method	prec.	F1-score	C-index
Cox PH	**0.851**	**0.798**	0.682
Random SF	0.787	0.694	**0.695**
Survival SVM	0.640	0.400	0.602

3.2.3 Survival distributions comparison

Along with the performance comparison it is important to review the predictions from the business point of view. It is arguable that the trend for the distribution of active customers should be decreasing with a longer lifespan. The Figure 3.3 provides a comparison between the distributions estimated with models described above. It is noticeable that RSF model better captures the decreasing trend and shares the similar behavior of estimated Kaplan-Meier survival curve (Figure 3.1). To verify the correctness of estimated predictions results were compared with historical median, standard deviation and average lifetime of customers.

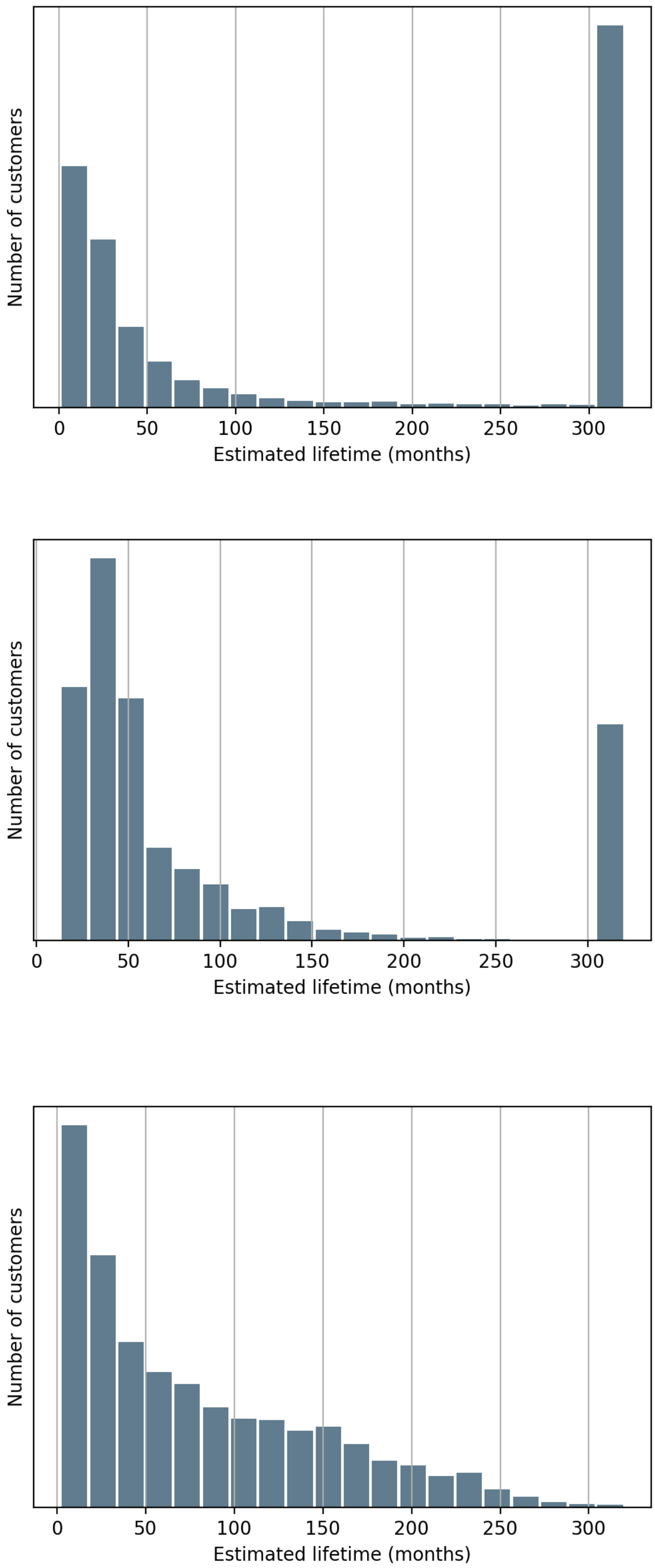

FIGURE 3.3: Distribution of number of customers for different lifespan duration. From top to bottom: CPH, Survival SVM, RSF

3.3 Behaviour transition model

Having the lifetime estimates, the next step is to predict customers behaviour. Since it is extremely complex to build the predictions on the customer level, the population could be divided into segments [1] based on the revenue behaviour. The assumption of this framework is that customer changes his or her revenue behaviour during the whole lifetime, therefore, the behaviour can be modeled as a transition between defined segments.

3.3.1 Customer segmentation

Customer segmentation is one of the central tasks in financial services. From the nature of those services follows that the companies cannot discriminate in terms of locality, region or other demographic specifics [25]. This means that variables such as sex, religion, social status, and other demographic features cannot impact the model predictions. Therefore, the model should rely on the behavior patterns and carefully follow the rules of ethics in AI [21]. The approach to identify market segments by casual factors instead of descriptive variables is called "benefit segmentation" [16]. The idea behind it is that customers can be divided into segments based on the benefits they are seeking in the products. Combining this with the revenue behavior provides us a detailed understanding of the customer and help us to build efficient segmentation engine. The revenue behavior is based on the historical values of RAR. Therefore, our segmentation engine relies not only on the potential profit but addresses risk measurements. We define these segments as behavioral states and each of them affects the revenue in a different manner.

3.3.2 Markov Model

A transition between the defined segments follows Markov process, therefore, the behavior changes can be modeled with Markov chain [28]. The probability of transition can be represented as a transition matrix, where (i, j) element with i^{th} row and j^{th} column is a number of customers who moved from segment i to segment j within the observed time period. Figure 3.4 and equation 3.5 represent the example of transition graph and the corresponding transition matrix T.

$$T = \begin{pmatrix} 0.7 & 0.2 & 0.1 \\ 0.3 & 0.6 & 0.1 \\ 0.6 & 0.2 & 0.2 \end{pmatrix} \tag{3.5}$$

Having the transition matrix, it is possible to estimate the probability of transitions during the whole lifetime, taking the matrix to the power of $i \in [1, n)$, where n is the duration of the lifetime. The disadvantage of this approach is that Markov chain model is stateless, so it does not take into account the previous behavioral history and additional covariates.

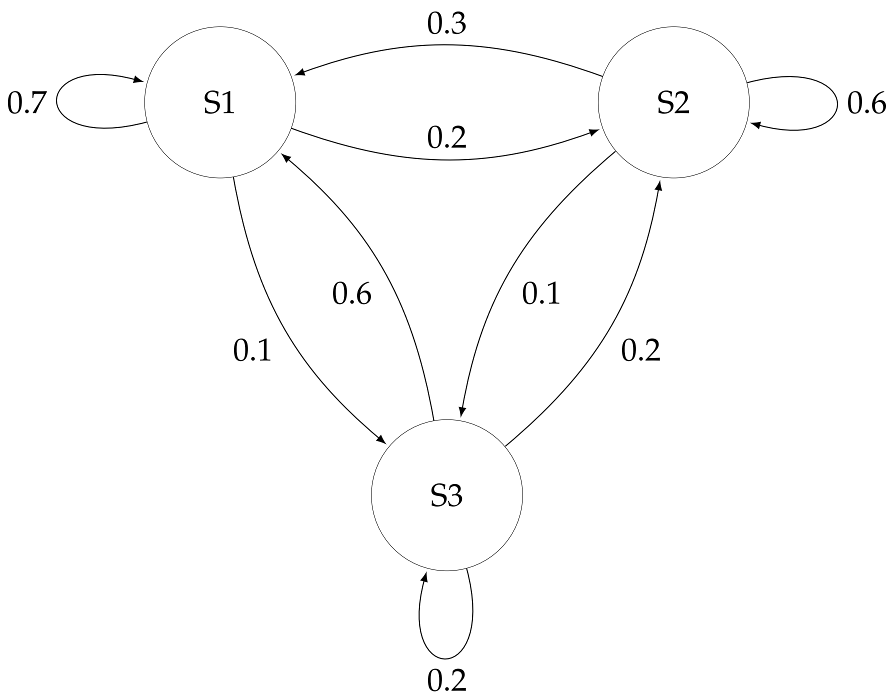

FIGURE 3.4: Example of transition graph between S1, S2 and S3 segments

3.3.3 Transition function approximation

Instead of having a huge matrix with all possible states and their combination, we can approximate function $P(S_n, C_n) = S_{n+1}$ where S_n is a vector of previous transitions and C_n vector of additional covariates. To get the prediction for the next time period we have to shift transitions history, so $S_n = S_{n+1}$. Nowadays, with recent progress in machine learning, there are multiple ways to approximate this kind of function, from the easily interpretable models as Logistic Regression and Decision Tree to a complex black-box model as a deep neural network.
Moreover, this method can handle a previously unseen sequence of transitions and provide customer behavior predictions, which cannot be implemented with Markov chain model.

Multinomial logistic regression

Multinomial logistic regression is a generalized version of a binary logistic regression model to distinguish between several classes. For this experiment was used one-vs-rest[2] learning algorithm. The separate model was trained for each class, so each binary model predicts weather objects belongs to a particular class. The next step is to apply softmax function to get a probability distribution for the next customer transition. This model is used as a baseline due to its high interpretability and simplicity.

Random Forest and XGBoost

Nowadays, tree ensemble methods became very popular in machine learning. It is a great tool for a nonlinear function approximation and suits perfect for the given problem. Random Forest [13] and XGBoost[5] algorithms were used to build a transition function.

3.3.4 Deep Neural Network

One of the main methods to address a problem with function approximation is a neural network. The assumption in the transition model is that customer behavior depends on his or her history of moving between states. Therefore, the goal of the model is to find a hidden relation between the states and their impact on the following one.
The recurrent neural networks are known to be good at capturing patterns when given a sequence data. Our network consists of three major parts: embedding layer, LSTM[14] layer and Dense layer.

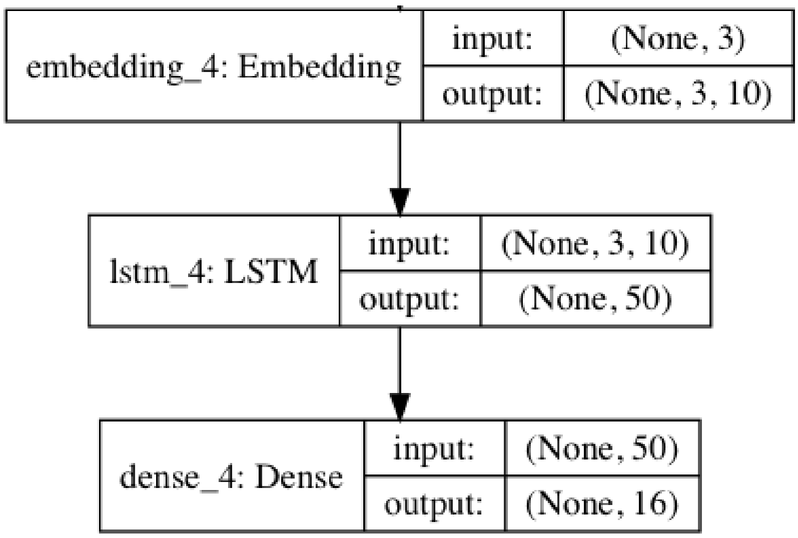

FIGURE 3.5: Architecture of DNN model

To map states, which are discrete categorical variables, to vectors, was used a technique called neural network embeddings[10]. The idea behind it is that each state can be represented as a vector of length n, therefore, some states can be closer or further in a transformed space.
Besides the similarity between states, the model aimed to find patterns in customer behavior changes. LSTM layer was used to approach this goal. It is known to be good in learning long-term sequences, so should handle the history of customer transitions.
Finally, the dense layer is used to map the learned patterns to the softmax function, to get a probability of transition to the each following state.

3.3.5 Evaluation and conclusion

The algorithms described above were compared by two metrics: Precision and F1-Score. To evaluate models we compared the predictions of the model and current ground truth cluster for each user. For Markov Model were used last 2 segments for each customer to build a transition matrix and the current one for evaluation.

TABLE 3.2: Precision, F1-Score and concordance index comparison for Cox's Proportional Hazard model, Random Survival Forest and Survival SVM in conjunction with CPH

Method	prec.	F1-score
Dummy Classifier	0.159	0.159
Markov Model	0.373	0.310
Logistic Regression	0.595	0.627
Random Forest	0.595	0.623
XGBoost	**0.599**	0.632
DNN	**0.599**	**0.633**

The experiments prove that the assumption about the significance of transitions history and additional covariates. Each classification method outperforms stateless Markov Model and random dummy classifier. As we can see from evaluation results, DNN model slightly outperforms the rest approaches, however, due to the limited transitions history performs almost identical to multinomial logistic regression. The increase of customer transitions history can enhance the deep neural network performance and predict more robust results.

3.4 Customer lifetime value calculation

The CLV is a sum of the current value (CV) and future value (FV). For each segment, we predict the future revenue throughout the maximum possible lifespan. With estimated tenure for each customer, we predict the transition probabilities for next time periods. The profit for each segment π_{s_i} is multiplied by a probability P of transition to this segment at time period t and summed throughout the whole lifetime N. A discount rate σ is applied to derive a present value of future cash flows for each customer.

$$CLV = CV + FV = \sum_{0}^{N} \frac{\pi_s * P_t}{(1+\sigma)^t} \tag{3.6}$$

To simplify the explanation, the example provided below shows the calculation of CLV for the four segments, for a customer with an estimated tenure of 3-time units and the current value of 115 units. For this example the discount rate is assigned to 8%.

TABLE 3.3: Transition probabilities for t = 1

Segment	1	2	3	4
Transition Probability	0.6	0.2	0.1	0.1
Expected revenue	120	135	80	40

TABLE 3.4: table

Future value: 111.0, Discounted: 102.7

TABLE 3.5: Transition probabilities for t = 2

Segment	1	2	3	4
Transition Probability	0.4	0.4	0.1	0.1
Expected revenue	125	138	75	45

TABLE 3.6: table

Future value: 117.2, Discounted: 100.5

TABLE 3.7: Transition probabilities for t = 3

Segment	1	2	3	4
Transition Probability	0.3	0.6	0.15	0.05
Expected revenue	130	145	77	46

TABLE 3.8: table

Future value: 139.85, Discounted: 111.01

Summing current and future values we get CLV of 429.21 units. In terms of our framework, we calculate the CLV estimates for each customer and re-estimate it after some time period. This leads us to the next step of a decision making: how to properly allocate resources taking into account CLV estimates.

Chapter 4

Resources allocation

The main goal for the bank is to build a strategy to properly operate on credit limits to optimize portfolio by shifting the balance to most profitable customers. The current value of customers and estimates of future value provide stakeholders with a broader view and ability to plan resources allocation for a long period properly.
Based on the CLV metrics all customers can be divided into four groups: currently low and potentially highly profitable, currently high and potentially high, currently low and potentially high and currently low and potentially low.
Each group requires different treatment to maximize the profit and retention of customers and minimizes the risk. We provide a set of actions to take, which impact on credit portfolio management. Taking into account the future profitability of customers, we can effectively operate on credit limits and apply cross-selling and up-selling techniques. Based on the findings we can build effective credit limit increase and decrease strategy to maximize the customer lifetime value [6]. Furthermore, estimates of future customers value also impact on the marketing strategy. Therefore, we can invest in more profitable customer segments and manage marketing campaigns more effectively [4] by selecting the right audience and evaluate the performance of a particular campaign [12]. The results of marketing campaigns have a long-term effect; therefore, we can observe how marketing activities enhance customer-firm relationships and improve retention in a long time period. The strategy is described in Figure 4.1.
The CLV approach is salable and could be applied to other banking domains. Creating a mix of CLV metrics and specific domain features, we can seek new business opportunities:

- Credit limit optimization
- Credit automation
- Pricing and attrition
- Marketing campaigns

The CLV is an essential metric for business strategy decisions [19]: managing sales forces to allocate resources by focusing on customers with high customer lifetime value and managing cross-selling targeting campaigns by selecting appropriate customers and decide on pricing strategy. Another benefit of CLV metrics is early warnings signs. Customer lifetime value is used to detect defection rates on early stages by defining in which customer segments problem originates. Finally, CLV is a good indicator of return on investment (ROI). It is arguable that the objective of any company is to maximize the profit and keep the customers who generate high revenue. Customer lifetime value helps to determine the amount of resources a company wants to spend on retention of a particular customer to maximize ROI.

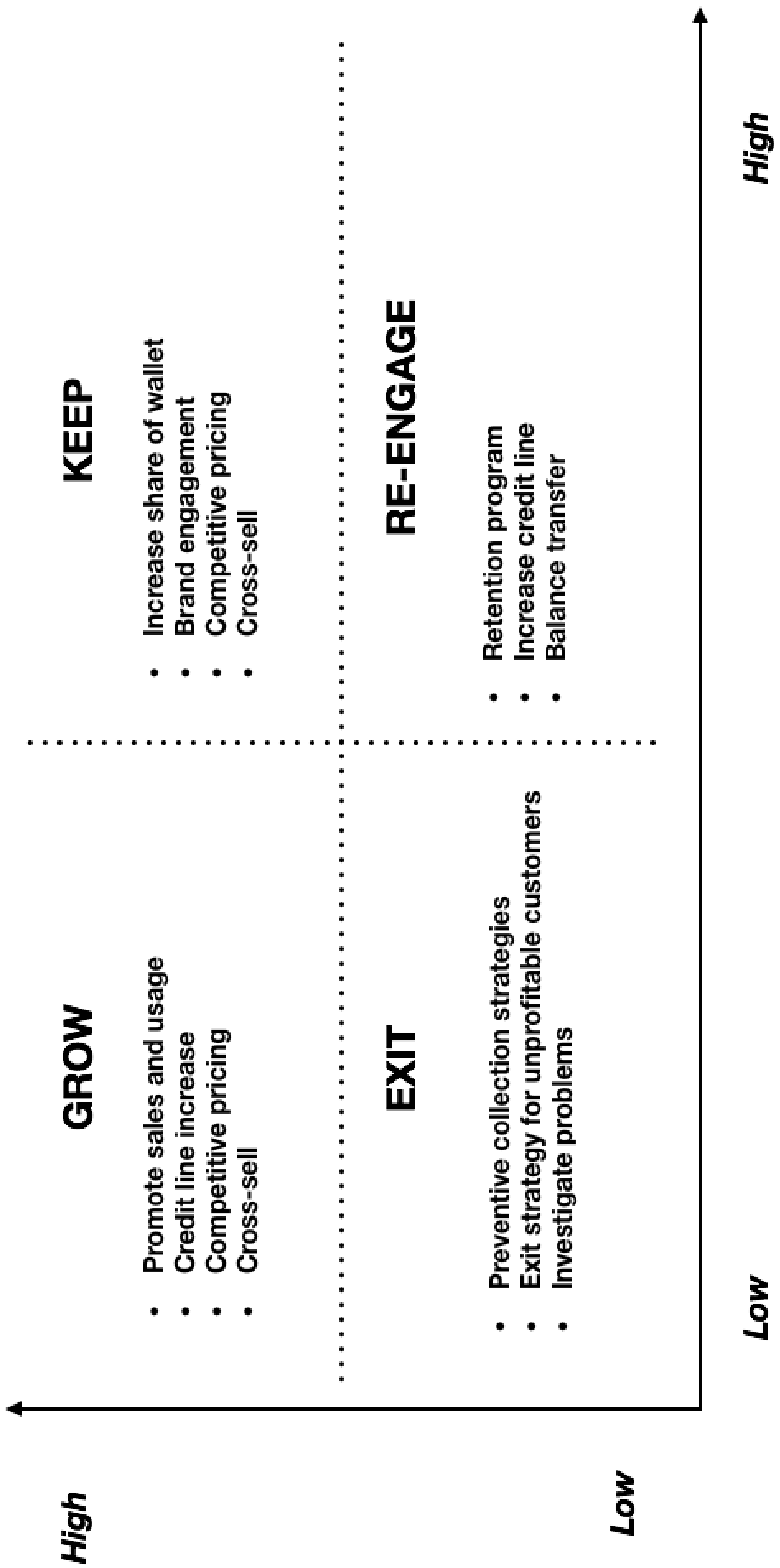

FIGURE 4.1: Resources allocation strategy based on the current and future values

Chapter 5

Conclusion

CLV is an important metric for banks to optimize a credit limit, improve retention and set competitive pricing. Vanilla CLV models estimate future value as a discounted cash flow over the lifetime of a customer; however, in credit cards market it is important to take into account a trade-off between revenue and risk. It is known that high-risk customers can generate a high revenue, but in the same time, such an aggressive strategy can lead to high losses.

In this study, we described a methodology for Customer Lifetime Value estimates in conjunction with a risk-adjusted return framework, based on the lifetime estimates, customer segmentation and behavior modeling.

In addition to this, we proposed a strategy for credit limit optimization and customer retention. We propose additional banking domains, where Customer Lifetime Value could be used in conjunction with specific domain features.

There are few limitations of our study that need to be acknowledged. Our methodology is developed only for existing customers and focuses only on revolving products. The framework has been tested on the credit cards market but can be scaled to credit lines as well. Unlike credit cards, lines of credit are usually secured, therefore, require a different risk treatment. The other limitation is that our framework does not address the costs of customer retention and maintenance.

The metric described in this work does not provide any monetary estimates of a credit limit for a particular customer, but provides recommendations on CLI/CLD strategy which can be used by stakeholders. The final limit heavily depends on the allowable level of risk and market specifics.

Bibliography

[1] Harsha Aeron et al. "A metric for customer lifetime value of credit card customers". In: *The Journal of Database Marketing Customer Strategy Management* 15 (Sept. 2008). DOI: 10.1057/dbm.2008.13.

[2] M Aly. "Survey on multiclass classification methods". In: *Survey on Multiclass Classification Methods* (Jan. 2005).

[3] Risk Management Group of the Basel Committee on Banking Supervision. "Principles for the management of credit risk". In: 1999. URL: https://www.bis.org/publ/bcbs75.pdf.

[4] Paul Berger and Nada I. Nasr. "Customer Lifetime Value: Marketing Models and Applications". In: *Journal of Interactive Marketing* 12 (Dec. 1998), pp. 17 –30. DOI: 10.1002/(SICI)1520-6653(199824)12:1<17::AID-DIR3>3.0.CO;2-K.

[5] Tianqi Chen and Carlos Guestrin. "XGBoost: A Scalable Tree Boosting System". In: (2016). DOI: 10.1145/2939672.2939785. eprint: arXiv:1603.02754.

[6] P. Clark. "Optimizing credit limit policies to maximise customer lifetime value". In: (2010).

[7] David Cox. "Regression Models and Life Table". In: *Journal of the Royal Statistical Society. Series B* 34 (Jan. 1972). DOI: 10.1007/978-1-4612-4380-9_37.

[8] Donald G Morrison David C Schmittlein and Richard Colombo. "Counting your customers: Who-are they and what will they do next?" In: Journal Management Science, 1987, pp. 1–24.

[9] Cameron Davidson-Pilon et al. *CamDavidsonPilon/lifelines: v0.17.0*. Jan. 2019. DOI: 10.5281/zenodo.2538195. URL: https://doi.org/10.5281/zenodo.2538195.

[10] Cheng Guo and Felix Berkhahn. "Entity Embeddings of Categorical Variables". In: (Apr. 2016).

[11] Sunil Gupta et al. "Modeling Customer Lifetime Value". In: *Journal of Service Research - J SERV RES* 9 (Nov. 2006), pp. 139–155. DOI: 10.1177/1094670506293810.

[12] Sunil Gupta et al. "Modeling Customer Lifetime Value". In: *Journal of Service Research - J SERV RES* 9 (Nov. 2006), pp. 139–155. DOI: 10.1177/1094670506293810.

[13] Tin Ho. "Random decision forests". In: vol. 1. Sept. 1995, 278 –282 vol.1. ISBN: 0-8186-7128-9. DOI: 10.1109/ICDAR.1995.598994.

[14] Sepp Hochreiter and Jürgen Schmidhuber. "Long Short-term Memory". In: *Neural computation* 9 (Dec. 1997), pp. 1735–80. DOI: 10.1162/neco.1997.9.8.1735.

[15] Arthur Middleton Hughes. *Strategic Database Marketing*. McGraw-Hill Pub. Co., 2005. ISBN: 007145750X.

[16] Russell I. Haley. "Benefit segments: Backwards and forwards." In: *Journal of Advertising Research* 24 (Feb. 1984), pp. 19–25.

[17] Hemant Ishwaran et al. "Random survival forests". In: (2008). DOI: 10.1214/08-AOAS169. eprint: arXiv:0811.1645.

[18] L., E Kaplan, and Paul Meier. "Nonparametric Estimation From Incomplete Observations". In: *Journal of American Statistical Association* 53 (June 1958), pp. 457–481. DOI: 10.1080/01621459.1958.10501452.

[19] Meko M.C. So and Lyn Thomas. "The 16 business benefits of customer lifetime value". In: (Nov. 2007).

[20] Fen Miao et al. "Is Random Survival Forest an Alternative to Cox Proportional Model on Predicting Cardiovascular Disease?" In: *IFMBE Proceedings* 45 (Jan. 2015), pp. 740–743. DOI: 10.1007/978-3-319-11128-5_184.

[21] DJ Patil Mike Loukides Hilary Mason. "Ethics and Data Science". In: (July 2018).

[22] Ka Lok Lee Peter S. Fader Bruce G.S. Hardie. "RFM and CLV: Using Iso-Value Curves for Customer Base Analysis". In: Journal Management Science, 2005, 415–430.

[23] Vikas Raykar et al. "On Ranking in Survival Analysis: Bounds on the Concordance Index". In: vol. 20. Jan. 2007.

[24] Lynette Ryals and Simon Knox. "Measuring risk-adjusted customer lifetime value and its impact on relationship marketing strategies and shareholder value". In: *European Journal of Marketing* 39 (May 2005), pp. 456–472. DOI: 10.1108/03090560510590665.

[25] Mohammad Safari et al. "Analyzing the Applications of Customer Lifetime Value (CLV) based on Benefit Segmentation for the Banking Sector". In: *Procedia - Social and Behavioral Sciences* 109 (Jan. 2014). DOI: 10.1016/j.sbspro.2013.12.511.

[26] David Schoenfeld, D.D.R. Cox, and D O. Oakes. "Analysis of Survival Data". In: *Journal of the American Statistical Association* 81 (June 1986), p. 572. DOI: 10.2307/2289259.

[27] Shweta Singh, B Murthi, and Erin Steffes. "Developing a measure of risk adjusted revenue (RAR) in credit cards market: Implications for customer relationship management". In: *European Journal of Operational Research* 224 (Jan. 2013), 425–434. DOI: 10.1016/j.ejor.2012.08.007.

[28] Robert Till and David Hand. "Behavioural models of credit card usage". In: *Journal of Applied Statistics* 30 (Feb. 2003), pp. 1201–1220. DOI: 10.1080/0266476032000107196.

[29] Vanya Van Belle and Sabine Huffel. "Support vector machines for survival analysis". In: *Proceedings of the Third International Conference on Computational Intelligence in Medicine and Healthcare* (Jan. 2007).

[30] R Van Oirbeek and Emmanuel Lesaffre. "Assessing the predictive ability of a multilevel binary regression model". In: *Computational Statistics Data Analysis* 56 (June 2012). DOI: 10.1016/j.csda.2011.11.023.

[31] Brandon Weathers and Richard Dr Cutler. "Comparision of Survival Curves Between Cox Proportional Hazards, Random Forests, and Conditional Inference Forests in Survival Analysis". In: 2017.