Aalto University

MS-E2177
Seminar on Case Studies in Operations Research

Project plan

Otto Saikkonen

Oscar Björklund

Aki Malinen

Iiro Sallinen

Juho Heimonen

March 2, 2020

1 Background

Our client is Nordea Bank, a full-service universal bank, with a total operating income of EUR 8.6 billion and total assets of EUR 554.8 billion in 2019 [1]. In our case study, we will focus on Nordea's insurance division, Nordea Life & Pensions, which aims to be the leading provider of life and pension products in the Nordic countries. The subsidiary is tightly integrated with the rest of the Nordea Group, to ensure seamless offers to Nordea's customers in all key life events [3]

For the insurance industry, the premiums of customers are the main profit sources while the settlement of claims and customer acquisition costs are the major expenditures. The insurance industry has the highest customer acquisition costs of any industry. It costs seven to nine times more for an insurance agency to attract a new customer than to retain one.[2] The premium margin, customer relation length and the claim risk are thus the most important aspects for insurance companies for evaluating customers [4]. Therefore, the best customers for insurance companies are the customers with high profits, long customer relationships and low risks.

One way to measure customer value has been to customer lifelong profitability, also known as customer lifetime value (CLV). In 1999, Mulhern proposed a computation model, where CLV was calculated as sum of discounted net cash flows brought in by the customer [4]. Especially with life insurance products, risks (death) are hard to predict. The main focus of our project will be put to analysis and prediction of incoming positive cash flows.

2 Objectives

The goal of this project is to develop a model for predicting the customer lifetime value (CLV) of a risk insurance customer. Ideally this model can be to be used to set standards for targeted marketing. The measure should be applicable in the context of life insurance industry. After we have defined a suitable measure for CLV, we aim to analyze the typical features that are related to different values of customer lifetime value. Furthermore, we aim to construct a predictive model that is capable of predicting the potential CLV from customer features that do not indicate the CLV explicitly. The model have to me validated using customer data that is not used in the fitting process.

The motivation of calculating CLV and providing estimation of potential CLV is the following: better diagnostics for these would enhance the ability to target potentially valuable customers. While Nordea Bank has hundreds of thousands of bank customers in Finland, it would be crucial to recognize the potentially most valuable customers from the point of view of life insurance sector. There is only limited amount of resources that can be used to actively contacting potential customers and it is important that the contacted customers would be profitable and also interested in the product. In the context of this project, we

focus on determining the potential CLV for customers and mostly leave out the estimation of which customers would were the most eager to buy.

Furthermore, constructing a suitable measure for the CLV would enable overall analysis of the current customers. This would help to evaluate the validity of common business claims such as that "20% of customer produce 80% of the profit".

3 Tasks

The main objective is to find a good metric for customer lifetime value in the context of life insurance customers like mentioned before. Second priority is determining monetarily best groups to advertise these insurance products to.

Firstly we need to see what data we need in order to determine an accurate CLV metric. This includes dialog with Nordea to get the attributes we are interested in. After the data delivery the main task is to analyse the data and confirm that the fore mentioned attributes are of sufficient quality and form.

The basis of CLV is the present value of the cash flow of the customer. This poses a clear problem in that the churn rates and realisation dates of the insurances are unknown. The task is therefore to make a model that predicts when a given customers cash flow will stop whether it be because of death or otherwise. More specifically we aim to find the expected value of cash flows for groups of customers or how some attributes contribute to the value of a customer e.g. age, sex and location of redidence. Ideally churn and death rates would be taken into account as separate events.

The model needs to be trained as well as validated with the data set from Nordea. After a functioning clv metric we can aim to solve the secondary objective. In the best case scenario this could also be done with the data set but more likely we will need additional information or support from Nordea. Alternatively the marketing team in Nordea could work independently with the CLV results.

4 Schedule

The project is divided into four overlapping work packages:

- Project planning and background research
- Project initiation
- Technical Execution
- Project management

The work package related to project planning & background research mostly consists of literature reviews and it is focused in the January - February period.

Project Initiation phase started when we received the initial description of the data structures available for our use. The technical execution phase starts once we receive a complete sample of the real project data. The project management work package overlaps with all of the other packages from beginning to end.

Complete GANTT chart and are schedule is available the Appendix.

5 Resources

Our team consists of five students in Mathematics and Operations Research, all with good knowledge of mathematical and statistical methods that are useful in this project. However, our knowledge of the insurance industry is inadequate, so we will allocate some time to study the fundamentals to be able to apply the mathematical methods for the problem. In our team, we are ensuring to distribute the work equally between the team members according to our individual strengths. Aki Malinen will be the manager of this project.

Our contact at Nordea, Almo Jaluta, will provide assistance during the whole project, as well as relevant data and material. Almo have already provided great ideas for how a CLV model can be implemented, and will be a great resource throughout the project. Further, the whole data analytics team at Nordea Life & Pensions support us throughout the project.

In addition to data and material provided by the client, we will study literature on CLV modelling and investigate what kind of methods have been tried and how they have performed. We aim to use academic research as our main source. Apart from our project manager Aki, we also have great support from both Professor Ahti Salo (Aalto University) and course assistant Roni Sihvonen (Aalto University).

To implement the solution, the team is allowed to use Python or R.

6 Risks

Risk	Probability	Effect	Impact	Mitigation Strategy
Poor data quality	High	Decrease quality of final solution	High	Careful inspection of data, filter and clean data.
Model too complex for the scope of the course	Medium	Too wide problem to solve for the allocated time	High	Clear formulation and scoping of the problem. Establish explicit goals.
Data security	Low	NDA contract violation	High	Local data management, Risk assessment preceding deadlines
Insufficient communication between team members	Medium	Resentment due to imbalance in workload between team members, misunderstandings	Medium	Regular communication between team members and manager and scheduling
Insufficient communication between team and client	Medium	Client not satisfied with the solution	High	Regular communication with the client
Team member inactivity or dropout	Low	High workload for other team members	High	Good communication between the project manager and the rest of the team. Clear schedule.

References

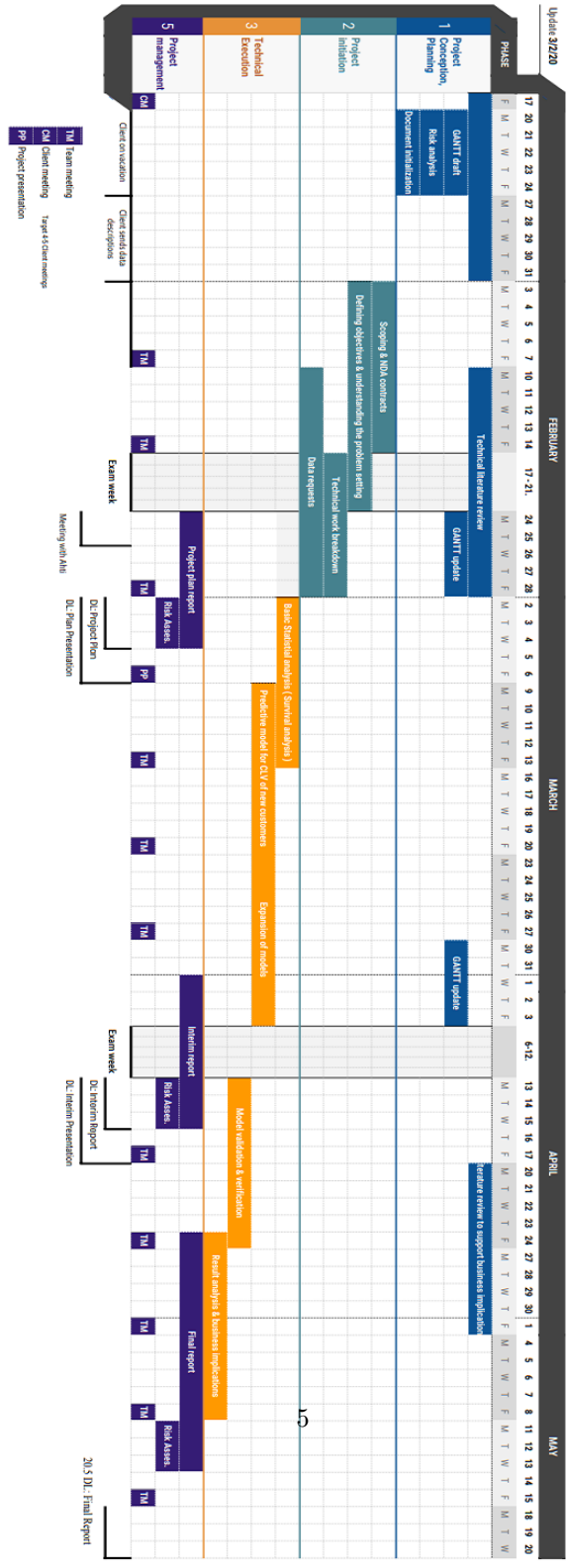
[1] *Nordea at a glance*,
`https://www.nordea.com/en/about-nordea/who-we-are/nordea-at-a-glance/`

[2] *Customer Loyalty And Retention Primer, cited 2.3.2020*,
Lynn Thomas
`https://www.iiadallas.org/page/75`

[3] *Fourth quarter and full year results 2019*,
Nordea

[4] *Customer profitability forecasting using Big Data analytics: A case study of the insurance industry*,
K. Fang, Y. Jiang b, M. Song
Computers & Industrial Engineering 101 (2016) 554–564

Appendices


GANTT: Predicting lifetime value for decision-making strategy development
Update 3/2/20
FEBRUARY
MARCH
APRIL
MAY
PHASE
1 Project Conception, Planning
2 Project Initiation
3 Technical Execution
5 Project management
GANTT draft
Risk analysis
Document initialization
Scoping & NDA contracts
Defining objectives & understanding the problem setting
Technical work breakdown
Data requests
Technical literature review
GANTT update
Basic Statistical analysis (Survival analysis)
Predictive model for CLV of new customers
Expansion of models
Model validation & verification
Result analysis & business implications
Final report
Literature review to support business implication
Project plan report
Risk Asses.
Interim report
Risk Asses.
Final report
Risk Asses.
Exam week
Exam week
Client on vacation
Client sends data descriptions
Meeting with Alisi
DL: Project Plan
DL: Plan Presentation
DL: Interim Report
DL: Interim Presentation
20.5 DL: Final Report
TM Team meeting
CM Client meeting
PP Project presentation
Target 4-5 Client meetings