



**HAL**  
open science

## Including Customer Lifetime Value in tree-based lapse management strategy

Mathias Valla, Xavier Milhaud, Anani Ayodélé Olympio

► **To cite this version:**

Mathias Valla, Xavier Milhaud, Anani Ayodélé Olympio. Including Customer Lifetime Value in tree-based lapse management strategy. 2022. hal-03903047v1

**HAL Id: hal-03903047**

**<https://hal.science/hal-03903047v1>**

Preprint submitted on 16 Dec 2022 (v1), last revised 12 Sep 2023 (v4)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Including Customer Lifetime Value in tree-based lapse management strategy

Mathias VALLA<sup>a</sup>, Xavier MILHAUD<sup>b</sup> and Anani OLYMPIO<sup>a</sup>

<sup>a</sup>LSAF, 50 av. Tony Garnier, Lyon, 69007, France

<sup>b</sup>I2M, Fred Joliot Curie, Marseille, 13001, France

## ARTICLE INFO

### Keywords:

Customer lifetime value  
Machine Learning  
Lapse  
Lapse management strategy  
Tree-based models  
Competing risks

## ABSTRACT

A retention strategy based on an enlightened lapse modelization can be a powerful profitability lever for a life insurer. Some machine learning models are excellent at predicting lapse, but from the insurer's perspective, predicting which policyholder is likely to lapse is not enough to design a retention strategy. Changing the classical classification problem to a regression one with an appropriate validation metric based on Customer Lifetime Value (CLV) has recently been proposed. In our paper, we suggest several improvements and apply them to a sizeable real-world life insurance dataset.

We include the risk of death in the study through competing risk considerations in parametric and tree-based models and show that further individualization of the existing approach leads to increased performance. We show that survival tree-based models can outperform parametric approaches and that the actuarial literature can significantly benefit from them. Then, we compare how this framework leads to increased predicted gains for the insurer regardless of the retention strategy. Finally, we discuss the benefits of our modelization in terms of commercial and strategic decision-making for a life insurer.

## 1. Introduction

In life insurance, « lapse risk » or « persistency risk » is the risk that the policyholder will cancel the contract at a time other than when the issuer expected when pricing the contract (KPMG (2020)). This risk is not considered an insurance risk because the payment to the policyholder "is not contingent on an uncertain future event that adversely affects the policyholder". However, lapse management is still undoubtedly a primary concern for life insurers. Lapses may substantially affect a company's solvency, its future profits and cash flows (Buchardt (2014); Buchardt et al. (2015)) or its Asset and Liabilities Management (ALM) (Kim (2005); Gatzert and Schmeiser (2008); Eling and Kochanski (2013); Eling and Kiesenbauer (2014)). The importance of measuring lapse and churn behaviours is global; it goes from yielding individual estimations of the Customer Lifetime Values (CLV) to being an estimator of a firm's profitability (Gupta and Lehmann (2006); Gupta (2009)) or strength (Ascarza et al. (2018)). Therefore for a life insurer, an enlightened lapse management strategy (LMS) is critical for successful monitoring and steering.

Part of the literature on lapse management adopts an economic-centred point of view (Dar and Dodds (1989); Kuo et al. (2003); Kagraoka (2005); Cox and Lin; Kiesenbauer (2012); Russell et al. (2013); Sirak (2015); Vasudev et al.; Nolte and Schneider (2017); Poufinas and Michaelide (2018); Yu et al. (2019)); we refer the reader to the complete bibliometric

analysis on this topic by Shamsuddin et al. (2022) for a summarized view of all these references. This economic-centred research aims to determine lapse factors like interest rates, gross domestic product, or unemployment rates. They are driven by economic hypotheses such as the emergency fund hypothesis (lapsing is a way of constituting an emergency fund), the policy replacement hypothesis (lapsing will occur when one changes its policy) or the interest rate hypothesis (lapsing depends strongly on rate change and arbitration). On the other hand, a large part of the literature investigates the individual determinants of lapse with policyholder-centred approaches. Micro-oriented features such as policyholder's personal information or the policy characteristics have shown to give valuable insights into lapse behaviour (Renshaw and Haberman (1986); Milhaud et al. (2011); Eling and Kiesenbauer (2014); Hwang et al. (2022)). Ćurak et al. (2015) as well as Gemmo and Götz (2016)'s works indicate that policyholders' features such as age and the number of beneficiaries are significant lapse factors, whereas Sirak (2015) dismissed those results. More recently, Loisel et al. (2021) proposed a comparison of lapse management strategies based on an innovative evaluation metric derived from the Customer Lifetime Value (CLV). Hu et al. (2021) investigated the benefits of incorporating spatial analysis in lapse modelling, and Azzone et al. (2022) showed with an approach based on random forests that microeconomic features such as the company's commercial approach for instance - is determining in the lapse decision. In contrast, macro-economical features only have a limited effect. This variety of results – sometimes contradicting each other – demonstrates the active interest in this research problem.

This paper focuses on lapse management strategy and retention targeting. As in Ascarza et al. (2018); Loisel et al. (2021), our goal is not only to model the lapse behaviour

 mathias.valla@univ-lyon1.fr (M. VALLA);

xavier.milhaud@univ-amu.fr (X. MILHAUD); anani.olympio@cnr.fr (A. OLYMPIO)

 <https://mathias-valla.com/> (M. VALLA);

<http://xaviermilhaud.fr> (X. MILHAUD)

ORCID(s): 0000-0003-4760-7849 (M. VALLA); 0000-0002-3962-9434 (X. MILHAUD)

but rather to select a retention strategy that would create more profitable policyholders whether the strategy targets them or not. This work shows that a well-chosen strategy, based on individualized CLV and directed toward a well-chosen target, increases the insurer's expected profitability. A critical concept that motivates many CLV-driven decisions is that customers should be judged as assets based on their future profitability for the insurer. Thus, since retention often serves as the basis for CLV modelizations (Gupta et al. (2006); Donkers et al. (2007); Lemmens and Gupta (2020) - sometimes specifically designed for targeting tasks (von Mutius and Huchzermeier (2021)) - and since CLV considerations should drive retention management, it seems natural to extend the existing life insurance application linking those topics together. We make decision-making a central concern of our work and suggest proactive lapse management tools allowing the insurer to undertake actions to prevent the causes of lapse; that is opposed to a reactive management approach where decisions are taken after lapses and aim at recapturing lost policyholders.

The goal of this paper is to create an individualized CLV model that will be used to enhance classical binary churn models. We will then have a model for lapse management strategy and retention targeting that we further improve with tree-based survival analysis and competing risks considerations. The global framework is directly inspired by Loisel et al. (2021). We try in this paper to build from that existing work and bring our contributions and extensions to that topic without dismantling the building blocks it represents. We model an individual future CLV with a survival model for which the risks of death and lapse are treated as mutually exclusive competing risks. For this purpose, we introduce parametric approaches - Cox cause-specific and subdistribution models - as well as tree-based survival models - Random Survival Forest (RSF) and Gradient boosting survival analysis. We focus here on tree-based models as they can often be considered state-of-the-art models (Grinsztajn et al. (2022)). Thus we introduce tree-based machine learning algorithms for binary prediction, including Classification and Regression Tree (CART), Random forests (RF), and Extreme Gradient Boosting (XGBoost) to lapse behaviour modelling. CART and XGBoost (Milhaud et al. (2011); Loisel et al. (2021)) were used in the literature for lapse modelling but have yet to be applied to predicting life insurance lapses in a competing risk setting. To our knowledge, while Random Survival Forest has been used for churn prediction recently (Routh et al. (2021)), both RSF and Gradient boosting survival analysis have never been used for that purpose before in an actuarial context. Our contribution to the actuarial literature is twofold. First, we detail a two-step lapse management modelization: we fit parametric and tree-based competing risk individual survival models to estimate individualized future CLVs that are part of an evaluation metric for tree-based lapse management models. Second, this work includes a business-oriented discussion of the results achieved by this framework, which is missing

from existing similar approaches.

The results and discussions show that a CLV-based lapse management strategy very often outperforms a more classical binary classification approach, even with competing risks and individualized considerations. When the latter yields profitable retention gain, the former can produce higher profits, up to more than 60%. If a loss-inducing retention strategy is considered, our methodology limits the loss considerably, often setting 0 as a floor limit or even turning it into a profit-inducing retention strategy. Sensitivity analysis explores the influence of conjectural and structural parameters.

The rest of this paper is structured as follows. We briefly outline the data used in our study in Section 2. In Section 3, we then introduce the binary classification models we selected and detail our study's methodology, describing the classical and CLV-based performance measures and discussing substantial parametrization improvements over existing approaches. Then, Section 4 details our two-step methodology, with the parametric and non-parametric modelizations of individual survival predictions, in a competing risks framework and then their implementation in the tree-based classification approaches considered. Section 5 is dedicated to presenting the real-life application we considered and the different results it produces. Those results are analyzed and discussed in Section 6 with commercial and strategic decision-making orientations. Eventually, Section 7 concludes this paper.

## 2. Data

We apply our framework to a real-world insurance portfolio. For privacy reasons, all the data, statistics, product names and perimeters presented in this paper have been either anonymized or modified. All analyses, discussions and conclusions remain unchanged.

We illustrate our methodology with a life insurance portfolio from a french insurer contracted between 1997 and 2018. Each record in the data set represents a unique policy for a unique policyholder. In the following sections, we will often refer to a unique pair of policy and policyholder by the term "subject". The dataset contains 251,325 rows with 248,737 unique policies and 235,076 unique policyholders. It means that some policies are shared between several policyholders and that one individual can detain several insurance policies. The dataset contains 43 covariates described in Table 1.

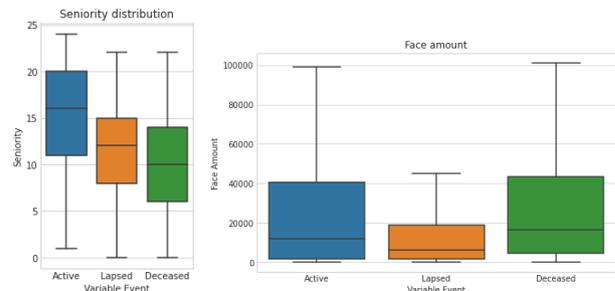
**Table 1**  
Data set description

	Covariates (Numerical, Categorical, Date)	Description
ID	CDI_ID_PERSONNE CDI_ID_CONTRAT	Policyholder (PH) unique ID Policy unique ID
PH-level informations	CDI_DT_NAISSANCE	PH's birth date (main PH when several poliholders owns one policy)
	Age_souscription	PH's age at subscription
	Nb_Contrats	Number of different policies owned by the policyholder
	CDI_CD_SEXE	PH's gender (1=Female; 2=Male; other=Non precised)
	CDI_DESTINATAIRE_COURRIER	Anonymised PH's name
	CDI_NUM_ET_NOM_VOIE	Anonymised PH's adress
	CDI_CD_POSTAL	Anonymised PH's post code
Policy-level informations	CDI_COMMUNE	Anonymised PH's city of residence
	CDI_TOP_ASSURE	Binary: 1 if PH is the main PH on the policy, 0 otherwise
	CDI_TYPE_PRODUIT	Type of product ("Top-end product" or "Classical product")
	CDI_NOM_PRODUIT	Name of life insurance product ("Product 1", "Product 2" or "Product 3")
	CDI_PARTENAIRE	Name of the insurance distributor
	CDI_DATE_DEB_CONTRAT	Policy's start date
	CDI_DATE_FIN_CONTRAT	Policy's end date
	START_YEAR	Policy's start year
	END_YEAR	Policy's end year
	SENIORITY	Policy's seniority (final seniority if the policy is ended, current seniority otherwise)
External data	STATE	Policy's state ("Active", "Lapsed", or "Death" if the policy ended following PH's death)
	YEAR	Last year of observation
Policy's cumulated financial flows	DISCOUNT RATE	Discount rate corresponding to the last year of observation
	TOTAL_PREMIUM_AMOUNT	Total face amount of the policy
	TOTAL_EURO_PREMIUM_AMOUNT	Face amount of the policy in euros
	TOTAL_UC_PREMIUM_AMOUNT	Face amount of the policy in units of account
	ARBITRATION_EURO	Cumulated arbitration amount of the policy in euros
	ARBITRATION_UC	Cumulated arbitration amount of the policy in units of account
	FEES_EURO	Cumulated fees amount of the policy in euros
	FEES_UC	Cumulated fees amount of the policy in units of account
	OTHER_EURO	Cumulated other part of the face amount of the policy in euros
	OTHER_UC	Cumulated other part of the face amount of the policy in units of account
	PREMIUM_EURO	Cumulated payments amount of the policy in euros
	PREMIUM_UC	Cumulated payments amount of the policy in units of account
	PROFIT_SHARING_EURO	Cumulated profit sharing amount of the policy in euros
	PROFIT_SHARING_UC	Cumulated profit sharing amount of the policy in units of account
	Covariates derived from financial flows	CLAIM_EURO
CLAIM_UC		Cumulated partial or total lapsed amount of the policy in units of account
%TOTAL_UC_PREMIUM_AMOUNT		Percentage of the face amount in units of account
%TOTAL_EURO_PREMIUM_AMOUNT		Percentage of the face amount in euros
%CLAIM_UC		Percentage of the face amount in units of account that was lapsed
Target covariate	%CLAIM_EURO	Percentage of the face amount in euros that was lapsed
	%CLAIM	Percentage of the total face amount that was lapsed
	EVENT	Policy's state (0=Active, 1=Lapsed, 2 ended following PH's death)

The data set represents policies that are majority owned by men (57.4%) for a mean censored seniority time of 13.4 years. Three products are present in the dataset. Product one was chosen by 72% of policyholders, product 2 by 25% and product 3 by 3%.

Regarding their state, 61% of the policies are still active, 22% lapsed, and 17% ended after the PH's death. We chose here to present the distribution of the variable *SENIORITY* as it is the response variable in our survival models. Its modelization has a critical influence on CLV, thus, on our lapse management strategy framework. We also chose to show the distribution of the variable *TOTAL PREMIUM AMOUNT* representing the most recent observed face amount for every subject, as it is a known determinant of lapse behaviour. We are aware that this covariate is a rather dynamic one as its value is updated at every payment, total or partial lapse, profit sharing, arbitration or even fees movements on a policy, and only considering its most recent value ignores a large part of the insights it can provide. Without any better option, we can only use *TOTAL PREMIUM AMOUNT* as it is and defer any dynamic considerations for future work.

The seniorities and most recent face amount recorded before the potential end of the policy are distributed as in Figure 1:



**Figure 1:** Seniorities and face amounts distributions

Without further analyzing the data, we can note several things. First, we can see that the mean censored seniority of 13.4 years is not equally distributed among our subjects. Active contracts tend to be older than lapsed ones, themselves older than policies that ended with the policyholder's death. That emphasizes the importance of several contributions, and the apparent difference in seniority regarding the cause of the policy's termination encourages a competing risks approach to analyze survival. Moreover, if we suspect that lapse and death are highly dependent on individual characteristics - such as the policyholder's age, for instance - this also supports an individualized survival analysis. Eventually, we can see that the last face amount observed is significantly lower for lapsed policies. It confirms our first intuitions, and the face amount will be included in our modelization.

Among the covariates introduced in Table 1, several ones play a central part in our two-step modelization: For the competing risks survival analysis step, *SENIORITY* will be the response variable, and all other covariates including individual data and financial flows are potential explanatory variables. The binary classification second step aims at predicting the *EVENT* outcome with minor transformations explained in Section 3 below. This prediction is based on the covariates underlined in Table 1.

### 3. Framework

The next section describes a modelization that follows Loisel et al. (2021)'s work. Our contributions place our work in a framework that differs from it in ways that will be made explicit, yet we chose to use a majority of their existing notations here. Usual lapse management models based on classification aim to predict whether a policyholder will lapse. They may perform very well at that specific task, but it only reflects some aspects of this economic problem. Indeed, we can easily imagine that many policyholders may be predicted as "lapsers" but may not be profitable to the insurance company. In that case, keeping such policyholders would be absurd, and an efficient model should not predict them as targets. Targeting policyholders is an economic problem that requires an economic measure to assess. We propose to consider a measure based on the discounted expected profit of all the policies, in other words, the sum of all (CLVs). Optimizing a lapse, churn or any other prediction task with business-related measures is not new. However, to our knowledge, none of the existing approaches uses individualized future CLVs and models the profit of retention strategy by accounting for competing risks or using survival tree-based models.

CLV is a well-studied subject in marketing and business economics in general and has also been modelled in an insurance context. For a given subject  $i$ , her future (CLV)

at horizon  $T$  can be modelled as follows:

$${}^F CLV_i(\mathbf{p}_i, \mathbf{F}_i, \mathbf{r}_i, \mathbf{d}, T) = \sum_{t=0}^T \frac{p_{i,t} F_{i,t} r_{i,t}}{(1 + d_t)^t}.$$

With  $t$  in years,  $t = 0$  represents the last observation point for subject  $i$ . The quantity  $p_{i,t}$  is her profitability ratio as a proportion of  $F_{i,t}$ , representing her face amount observed at time  $t$ . The quantity  $r_{i,t}$  is the  $i$ -th subject's probability of still being active at time  $t$  and naturally,  $d_t$  is the discount rate at time  $t$ , common to every subject. It is worth pointing out that  ${}^F CLV_i$  does not represent the global profit generated by subject  $i$  from her policy's first year until time  $T$ ; it represents the future  $T$  years of profit.

We suggest a modelization of the insurer's estimated profit - or loss - resulting from a lapse management strategy. In order to do that, we will compare the expected value of the portfolio before and after applying a given strategy. We are aware that there could be infinite ways to design a retention campaign: offering a punctual incentive, recurrent services or more profit sharing, for instance. Here, we define what we will consider a lapse management strategy (LMS):

**Definition 1 (Lapse management strategy (LMS)).**

A  $T$ -years lapse management strategy is modelled by the offer of an incentive  $\delta_i$  to subject  $i$  if she is targeted. The incentive is expressed as a percentage of her face amount and is accepted with probability  $\gamma_i$ . Contacting the targeted policyholder has a fixed cost  $c$ . A targeted subject who accepts the incentive will be considered as an "acceptant" who will never lapse, and her probability of being active at year  $t \in [0, T]$  is denoted  $r_{i,t}^{acceptant}$ . Conversely, a subject who refuses the incentive and prefers to lapse will be considered as a "lapser", and her probability of being active at year  $t$  is denoted  $r_{i,t}^{lapser}$ . A lapse management strategy is uniquely defined by the parameters  $(\mathbf{p}, \boldsymbol{\delta}, \boldsymbol{\gamma}, c, T)$

Even if this definition is already a simplification of any real-life insurance retention strategy, various constraints and the data and tools at the insurer's disposal do not always allow to conduct such a study. In the following section, we consider a simplified version of this framework by assuming that  $p_{i,t}$ ,  $F_{i,t}$ , and  $d_t$  remain constant across time, and denoted  $p_i$ ,  $F_i$  and  $d$  hereafter, with  $F_i$  being the most recent face amount observed for subject  $i$ . Moreover, we set  $\gamma_i$  and  $\delta_i$  to be the same for all subjects and denoted as  $\gamma$  and  $\delta$  hereafter. Finally, the last observed state of subject  $i$  is denoted  $y_i$ , with  $y_i = 1$  if the policy is lapsed,  $y_i = 0$  otherwise.

With those considerations, we can then define the control portfolio's future value as follows:

$${}^F CPV(p, \delta, \gamma, c, T) = \sum_{i=1}^n {}^F CLV_i(p_i, F_i, \mathbf{r}_i^{\text{acceptant}}, d, T) \cdot \mathbf{1}(y_i = 0) + \sum_{i=1}^n {}^F CLV_i(p_i, F_i, \mathbf{r}_i^{\text{lapsers}}, d, T) \cdot \mathbf{1}(y_i = 1)$$

It represents the hypothetical value of the portfolio, considering that:

- every subject that did not lapse up to her last observation point -  $y_i = 0$  at  $t = 0$  - has a vector of retention probabilities of  $\mathbf{r}_i^{\text{acceptant}}$ ;
- every subject that has been observed to lapse -  $y_i = 1$  at  $t = 0$  - has a vector of retention probabilities of  $\mathbf{r}_i^{\text{lapsers}}$

**Remark 1.** It is important to note that this does not reflect the actual future value of the portfolio - as the future CLV of lapsers should be 0 - but rather its hypothetical expected future value given the nature (lapsers or not) of every subject but not their actual states (actually lapsed or not). It represents this hypothetical future CLV of all subjects if no customer relationship management about lapses is carried out.

A classification algorithm would take the lapse indicator  $y_i$  as a target variable and yield predictions  $\hat{y}_i$ . Given a lapse management strategy and such a classification algorithm, we define the lapse managed portfolio future value by:

$${}^F LMPV(p, \delta, \gamma, c, T) = \sum_{i=1}^n {}^F CLV_i(p, F_i, \mathbf{r}_i^{\text{acceptant}}, d, T) \cdot \mathbf{1}(y_i = 0, \hat{y}_i = 0) + \sum_{i=1}^n {}^F CLV_i(p, F_i, \mathbf{r}_i^{\text{lapsers}}, d, T) \cdot \mathbf{1}(y_i = 1, \hat{y}_i = 0) + \sum_{i=1}^n {}^F CLV_i(p - \delta, F_i, \mathbf{r}_i^{\text{acceptant}}, d, T) \cdot \mathbf{1}(y_i = 0, \hat{y}_i = 1) + \gamma \cdot \sum_{i=1}^n {}^F CLV_i(p - \delta, F_i, \mathbf{r}_i^{\text{acceptant}}, d, T) \cdot \mathbf{1}(y_i = 1, \hat{y}_i = 1) + (1 - \gamma) \cdot \sum_{i=1}^n {}^F CLV_i(p, F_i, \mathbf{r}_i^{\text{lapsers}}, d, T) \cdot \mathbf{1}(y_i = 1, \hat{y}_i = 1) - \sum_{i=1}^n c \cdot \mathbf{1}(\hat{y}_i = 1)$$

Clearly, the sums appearing in the formulas above could be grouped to make them more concise. We chose not to do so for the sake of visualization: we can distinctly see each possible case in each summand.

Then, we define the economic metric of the algorithm as the retention gain, the future profit generated by the retention management strategy over  $T$  years:

$$RG(p, \delta, \gamma, c, T) = {}^F LMPV(p, \delta, \gamma, c, T) - {}^F CPV(p, \delta, \gamma, c, T)$$

that can be simplified as follows:

$$RG(p, \delta, \gamma, c, T) = \sum_{i=1}^n \left[ \gamma \left[ {}^F CLV_i(p - \delta, F_i, \mathbf{r}_i^{\text{acceptant}}, d, T) - {}^F CLV_i(p, F_i, \mathbf{r}_i^{\text{lapsers}}, d, T) \right] \cdot \mathbf{1}(y_i = 1, \hat{y}_i = 1) - {}^F CLV_i(\delta, F_i, \mathbf{r}_i^{\text{acceptant}}, d, T) \cdot \mathbf{1}(y_i = 0, \hat{y}_i = 1) \right] - \sum_{i=1}^n c \cdot \mathbf{1}(\hat{y}_i = 1)$$

This evaluation metric can now be derived into an individual retention gain measure. More specifically, we define  $z_i$  as:

$$z_i = \begin{cases} -{}^F CLV_i(\delta, F_i, \mathbf{r}_i^{\text{acceptant}}, d, T) - c & \text{if } y_i = 0 \\ \gamma \cdot \left[ {}^F CLV_i(p - \delta, F_i, \mathbf{r}_i^{\text{acceptant}}, d, T) - {}^F CLV_i(p, F_i, \mathbf{r}_i^{\text{lapsers}}, d, T) \right] - c & \text{if } y_i = 1 \end{cases}$$

That last equation can seem obscure. It only assigns to each individual the expected profit or loss that would result from targeting her with a given lapse management strategy. A positive amount for subject  $i$  means that targeting her would generate profit, whereas a negative one would lead to a loss for the insurer. We can take the example of a hypothetical scenario where  $p_i = 3\%$ ,  $\delta = 0.05\%$ ,  $\gamma = 10\%$  and  $c = 10$  euros. It would generate  $z_i$ s taking values from  $-234,614\text{€}$  to  $53,066\text{€}$  with a mean of  $-218\text{€}$  and a median of  $-55\text{€}$ . Different scenarios would result in very different distributions for the  $z_i$ 's.

Eventually, we define  $\tilde{y}_i$  as a binary target variable indicating for policyholder  $i$  if the individual expected retention gain resulting from a given retention strategy is a profit or a loss. More specifically, we define  $\tilde{y}_i$  as:

$$\tilde{y}_i = \begin{cases} 1 & \text{if } z_i > 0 \\ 0 & \text{if } z_i \leq 0 \end{cases},$$

**Remark 2.** A subject in the dataset for which  $y_i = 0$  would produce  $\tilde{y}_i = 0$ , whereas one for which  $y_i = 1$  could produce  $\tilde{y}_i = 0$  or  $\tilde{y}_i = 1$ . In other words, it is never profitable for the insurer to offer an incentive to a subject that would not have lapsed. On the other hand, offering that same incentive to a lapsers can be profitable. However, depending on the subject's features and the lapse management strategy parameters, it can also lead to a loss.

We can now include  $\tilde{y}_i$  as a new binary target variable in our models and directly consider  $RG$  as the global evaluation metric in the tree-based models we consider.

We can now compare two models: the classical one with  $y_i$  as a target variable and accuracy as the evaluation metric; and the CLV-augmented one with  $\tilde{y}_i$  as a target variable and  $RG$  as the evaluation metric.

Intuitively, the former tries to predict whether a policyholder will lapse and tune its parameters by minimizing the misclassification rate. On the other hand, the latter aims at predicting whether applying a given retention strategy to the  $i$ -th individual will be profitable for the insurer and tune its parameters by maximizing the global expected retention gain.

## 4. Methodology

In Section 3, we described a business-oriented framework, augmenting lapse management strategy with an evaluation metric based on the future CLV of every subject. Evaluating this metric requires computing  $r^{\text{acceptant}}$  and  $r^{\text{lapser}}$ , the matrices of size  $(n, T + 1)$  containing for every subject, survival probabilities that we detail below.

Given this framework, we propose a two-step methodology: firstly, we detail how this survival analysis is carried out to model those retention parameters, and secondly, we explain how we use them for training tree-based classification models.

### 4.1. Step 1: Modelling $r^{\text{acceptant}}$ and $r^{\text{lapser}}$

We recall that a given subject's policy can end with lapse or death, and the policy is considered active if competing events are yet to occur. Furthermore, while a lapser's policy can end with lapse or death, whatever comes first, an acceptant one can only end with death.

$r^{\text{lapser}}$  represents the probability that the policy of subject  $i$  is still active at time  $t$ , given that the subject is labeled as a lapser -  $\text{EVENT} = 1$  - at  $t = 0$ . Predicting these overall conditional survival probabilities with competing risks can be achieved by creating a combined outcome: the policy ends with death or lapse, whichever comes first. In order to compute  $r^{\text{lapser}}$  in practice, we recode the competing events as a combined event. This approach is compatible with any survival analysis method regarding statistical guarantees.

Conversely,  $r^{\text{acceptant}}$  represents the probability that the policy of subject  $i$  is still active at time  $t$ , given that the subject is not labelled as a lapser -  $\text{EVENT} = 0$  or  $2$  - at  $t = 0$ . This modelization is more complex as we must dissociate the risks of lapse and death. These causes being mutually exclusive, a competing risks methodology is well-suited to estimate  $r^{\text{acceptant}}$  (Laurent et al. (2016)).

#### 4.1.1. Competing risks frameworks

We are aware that the context of our modelization requires competing risk setting. As detailed in Appendix A.5, several regression models exist to estimate the global hazard and the hazard of one risk in such settings: cause-specific and subdistribution modelizations. They account for competing risks differently, obtaining different hazard functions and thus have distinct advantages, drawbacks and interpretations. These differences are discussed in Milhaud and Dutang (2018), where the authors also considered a competing risk framework for lapse prediction. After discussions detailed in Appendix A.5, the simplicity of

a cause-specific approach and the fact that it can be adapted to any survival method, including tree-based ones, oriented our choice towards it. We then computed  $r^{\text{acceptant}}$  and  $r^{\text{lapser}}$  with three different methods - Cox model, Random Survival Forest and Gradient Boosting Survival Model - and retained the best one. These methods are shortly described in the following sections.

#### 4.1.2. Cox proportional hazard model

One of the most common survival models is the Cox proportional hazard (CPH) model (Cox). It postulates that the hazard function can be modeled as the product of a time-dependent and a covariate-dependent functions. The hazard function at time  $t$  for subject  $i$  with covariate vector  $X_i$ , under Cox proportionnal hazard model can be expressed as:

$$\underbrace{\lambda(t|X_i^1, X_i^2, \dots)}_{\text{hazard function}} = \lambda(t|X_i) = \underbrace{\lambda_0(t)}_{\text{baseline hazard}} \underbrace{e^{(X_i \cdot \beta_i)}}_{\substack{\text{log-partial hazard} \\ \text{partial hazard}}}$$

It is crucial to note that in this model, the hazard function is the product of the baseline hazard, which only varies with time, and the partial hazard, which only varies depending on the covariates. The parameters of this model are the  $\beta$ , and they can easily be estimated with a maximum likelihood approach. Their estimation can be carried out without having to model  $\lambda_0(t)$  - which is why CPH is considered semi-parametric.

We use Python and lifelines (Davidson-Pilon (2019)) to implement it. We specify a spline estimation for the baseline hazard function. We select the covariates and model parameters using AIC (Akaike (1973)) and use the concordance index (Harrell et al. (1982)) to compare CPH to other models. The concordance index - or Harrel's c-index or simply c-index - is a metric to evaluate the predictions made by a survival model. It can be interpreted as a generalization of the area under a receiver operating characteristic (ROC) curve (Hanley and McNeil (1982)) - or AUC - in a survival setting with censored data.

#### 4.1.3. Random Survival Forest

Survival trees have been extensively studied for a long time, and a complete review of such existing methods up to 2011 can be found in Bou-Hamad et al. (2011). The most important thing to understand is that a survival tree can be created by modifying the splitting criterion of a regular tree. Most survival tree algorithms are designed with a split function that aims to maximize the separation of the resulting child nodes in terms of survival profiles. This separation between nodes is estimated by maximizing the log-rank statistic (Mantel (1966); LeBlanc and Crowley (1993)). Each terminal node of a survival tree contains a survival profile from which we can derive the survival and cumulative hazard function.

An RSF is the counterpart of a random forest (see Appendix A.6.2) for such survival trees. It has been developed in Ishwaran et al. (2008) and extended for competing risks a few years after (Ishwaran et al.). A prediction with RSF for a given subject is made by getting his/her survival profile in each tree in the forest. His/her corresponding survival and cumulative hazard function are estimated in each tree with Kaplan-Meier and Nelson-Aalen estimators, respectively. Eventually, the aggregation of those single-tree estimates constitutes the RSF's prediction.

We use Python and `sksurv` (Pölsterl (2020)) to implement RSF, and we tune and evaluate our model using the concordance index.

**Remark 3.** `sksurv` allows us to use RSF with a cause-specific consideration of the competing risks. To this day, `sksurv` does not have a subdistribution competing risks model, whereas its R implementation `randomForestSRC` does (Ishwaran and Kogalur (2007)).

Moreover, a severe limitation of that approach is that predictions can only be made at time points observed in the training set. Concretely, this prevents us from using RSF to extrapolate survival and hazard functions to unobserved time points.

#### 4.1.4. Gradient Boosting Survival Model

In the same way Random Forest has a survival counterpart, this is also true for Gradient Boosting approaches. An essential distinction between classical boosting algorithms (see Appendix A.6.3) and Gradient Boosting Survival Model (GBSM) lies in its loss function. The loss function that we use with GBSM is the partial likelihood loss of a CPH model, and the optimization in such a model is achieved by maximizing a slightly modified log-partial likelihood function:

$$\arg \min_f \sum_{i=1}^n \delta_i \left[ f(\mathbf{X}_i) - \log \left( \sum_{j \in \mathcal{Z}_i} e^{f(\mathbf{X}_j)} \right) \right]$$

Where  $\delta_i$  is the event indicator and  $f(\mathbf{X}_i)$  is GBSM's prediction for the  $i$ -th subject, with a covariate vector  $\mathbf{X}_i$ .  $g_i$  is the tree leaf including subject  $i$ . Similarly to RSF, we use Python and `sksurv` (Pölsterl (2020)) to implement GBSM. We tune and evaluate our model using the concordance index. Remark 3 also applies here.

#### 4.1.5. Final modelization choice

Our analysis shows that, based on concordance index, RSF and GBSM both outperformed a semiparametric Cox model in our study case. Regarding interpretability, we note that the feature importance analysis is very similar between the three models. All the details about the final concordance index scores, covariates importance and various plot for further analysis are available in Appendix A.6.

In the following sections, we decide to retain **GBSM** for the modelization of  $r^{\text{acceptant}}$  and  $r^{\text{lapser}}$  as it has the best concordance index.

**Remark 4.** As this study aims to be business-oriented and favour real-life decision-making, it is crucial to note that the computation times for fitting these different models are very different and could potentially be a huge constraint for real operational deployment. Specific computation times differ greatly depending on various factors, such as the number of subjects or features considered, the computation power or parallelization ability at disposal, for instance. However, we can still give here an order of magnitude for those differences. If the tuning and fitting process for CPH can last a few tens of seconds, it lasts hours for RSF and tens of hours for GBSM.

## 4.2. Step 2: Classification tasks

Our work focuses on lapse management with tree-based models. It aims to answer the question: which policyholders would be worth targeting with a lapse management strategy to maximize the expected T-year profit for the insurer? We will consider a single tree built with Breiman's CART algorithm, Random Forest, XGBoost, and RE-EM trees. The following sections detail how those different approaches work. Those models will be compared on two different classification tasks; and tuned with two different evaluation metrics, a statistical metric and a business-related one.

On  $y_i$  First, we will use a classical lapse prediction framework to model the policyholder's behaviour. Each policyholder will be labelled as a lapser or a non-lapser with a binary outcome  $y_i$ . Our first batch of models will be trained with  $y_i$  as a response variable and produce predictions  $\hat{y}_i$ . Accuracy( $y, \hat{y}$ ), which is undoubtedly the most intuitive performance measure for binary classification, is defined as the proportion of correctly predicted observations over all observations. It is widely used for churn analysis and appears to be a satisfying performance measure for relatively balanced outcomes in binary classification problems. We will use it as an evaluation metric in a 10-fold cross-validation step for tuning our models.

We know that more complex statistical evaluation metrics can be used for binary classification, such as the  $F_\beta$  Score family, the AUC under the ROC, the Brier Score, or the lift curve. In this paper, and similarly to a vast part of the applied binary classification literature, we will only select, evaluate and discuss the models in the light of the metrics we introduced.

On  $\tilde{y}_i$  Secondly, we will use the CLV-Augmented lapse prediction framework, detailed in Section 3. Each policyholder will be labelled as a targeted lapser or a non-targeted policyholder with the binary outcome  $\tilde{y}_i$  and prediction for that outcome are denoted  $\hat{\tilde{y}}_i$ .

**Remark 5.** Note that whenever  $y_i = 0$ , we also have  $\tilde{y}_i = 0$ . In other words, if subject  $i$  does not intend to lapse, it is never worth proposing her an incentive: the subject will accept it with probability 1 and would not have lapsed.

On the other hand, when  $y_i = 1$ , it corresponds to either  $\tilde{y}_i = 1$  or  $\tilde{y}_i = 0$ . In other words, if subject  $i$  is labelled as a lapses, it does not necessarily mean it is worth targeting her. From the insurer's point of view, some policies are better off lapsed.  $\tilde{y}$  can be seen as a more detailed version of  $y_i$  as it carries not only behavioural information regarding lapse but also a profitability one.

We thus train a second batch of models with  $\tilde{y}_i$  as a response variable. We use *RG* as an evaluation metric in a 10-fold cross-validation step for tuning these models.

*Summary of our methodology:* First, we train a CART, RF and XGBoost models with  $y_i$  as a binary target variable and accuracy as a tuning evaluation metric.

Then we train them with  $\tilde{y}_i$  as a binary target variable and *RG* as a tuning evaluation metric.

Finally, we train and test all six models on different random samples of our dataset and keep track of the model's classification performance on all of them and for various retention strategies for comparison's sake.

The sections below briefly introduce the tree-based model we selected before displaying how they performed in various lapse management scenarios.

#### 4.2.1. CART

CART (*Classification And Regression Trees*) is an algorithm developed by Breiman et al. (1984) that consists of recursively partitioning the covariate space. It is a widespread, intuitive and flexible algorithm that handles regression and classification problems.

#### 4.2.2. Random forest

A natural idea to correct CART's instability and enhance its prediction accuracy is the aggregation of a significant number of single trees, each grown on different subsamples of the dataset. A random forest (RF by Breiman (2001)) is a tree-based bagging procedure where each tree is grown on randomly drawn observations and contains splits considering only randomly drawn covariates.

#### 4.2.3. XGBoost

Other tree-based approaches have been designed to reduce the instability of a single tree model. Model boosting is an adaptative technique, first developed by Freund et Shapire (Freund and Schapire (1996)), that does not rely on the aggregation of independent weaker models but rather on the aggregation of weak models built sequentially, one after the other. XGBoost (Chen and Guestrin (2016)) is a widespread and performant tree-boosting model that relies on a gradient boosting step and provides a very optimized

parallelized procedure. It is considered a state-of-the-art library for various prediction problems.

The interested reader can find more detailed explanations about CART, RF and XGBoost mechanisms in Appendix A.1. For these modelizations, we used Python and sklearn (Pedregosa et al. (2011)).

## 5. Real-life application

Based on the real life-insurance dataset at our disposal (described in Section 2), we use the survival model we selected and estimate  $r^{\text{acceptant}}$  and  $r^{\text{lapses}}$  for every individual. This allows us to compute the individual CLVs, RGs,  $z_i$ 's and  $\tilde{y}_i$ . We have already defined what a strategy is (see Definition 1), and we can thus apply our classification methodology to various retention strategies.

### 5.1. Considered lapse management strategies

The strategies considered are based on several criteria. First, we selected realistic strategy parameters and time horizons based on actual retention campaigns led by life insurers. Moreover, we chose to present strategies that illustrate the exhaustive list of conclusions and discussions that are carried out in the next section. Finally, we also incorporated strategies that are "obviously bad" in the sense that such strategies would necessarily lead to a loss for the insurer. Such extreme scenarios will supplement our discussions.

Results related to the 32 considered LMS are given in Section 5.2. Related comments on how to read these tables are given in Section 5.3.

Scenarios	p	$\delta$	$\gamma$	c	d	T
1	2.50%	0.04%	25%	10	1.50%	5
2	2.50%	0.04%	25%	100	1.50%	5
3	2.50%	0.04%	25%	100	1.50%	20
4	2.50%	0.04%	5%	10	1.50%	5
5	2.50%	0.04%	5%	100	1.50%	20
6	2.50%	0.10%	25%	10	1.50%	20
7	2.50%	0.10%	25%	100	1.50%	20
8	2.50%	0.10%	5%	100	1.50%	20
9	1.50%	0.04%	25%	10	1.50%	20
10	1.50%	0.04%	25%	100	1.50%	5
11	1.50%	0.04%	5%	10	1.50%	20
12	1.50%	0.04%	5%	100	1.50%	5
13	1.50%	0.20%	20%	10	1.50%	5
14	1.50%	0.10%	25%	100	1.50%	5
15	1.50%	0.10%	5%	10	1.50%	20
16	1.50%	0.10%	5%	100	1.50%	5
17	1.50%	0.10%	5%	100	1.50%	20
18	2.50%	0.08%	20%	10	1.50%	20
19	2.50%	0.08%	20%	100	1.50%	5
20	2.50%	0.08%	20%	100	1.50%	20
21	2.50%	0.08%	10%	10	1.50%	20
22	2.50%	0.08%	10%	100	1.50%	20
23	2.50%	0.20%	20%	100	1.50%	5
24	2.50%	0.20%	20%	100	1.50%	20
25	2.50%	0.20%	10%	10	1.50%	20
26	2.50%	0.20%	10%	100	1.50%	20
27	1.50%	0.08%	20%	10	1.50%	5
28	1.50%	0.08%	20%	100	1.50%	5
29	1.50%	0.08%	10%	100	1.50%	5
30	1.50%	0.20%	20%	100	1.50%	5
31	1.50%	0.20%	10%	10	1.50%	20
32	1.50%	0.20%	10%	100	1.50%	20

Table 2: All considered LMS

## 5.2. Numerical results

N°	time (s)	Model	% target diff	Accuracy		Retention gain		RG/target		Improvement
				$y_i$	$\tilde{y}_i$	$y_i$	$\tilde{y}_i$	$y_i$	$\tilde{y}_i$	
1	4949	CART	62.6%	92.3%	85.3%	114 661	219 655	4.48	38.20	91.6%
		RF		92.9%	85.4%	232 314	287 884	9.82	56.65	23.90%
		XGB		93.4%	85.8%	243 365	324 952	9.61	54.64	33.50%
2	4603	CART	93.5%	92.3%	83.3%	- 2 187 622	- 8 224	- 85.52	- 31.09	99.60%
		RF		92.9%	83.4%	- 1 900 265	45 483	- 80.18	194.35	102.40%
		XGB		93.4%	83.5%	- 2 032 650	77 481	- 80.39	174.44	103.80%
3	5555	CART	55.4%	92.3%	86.5%	4 789 814	5 117 844	187.00	577.74	6.80%
		RF		92.9%	86.4%	4 463 796	4 255 175	188.47	566.05	-4.70%
		XGB		93.4%	86.8%	5 032 706	5 433 366	198.92	610.26	8.00%
4	4753	CART	86.7%	92.3%	83.6%	- 514 477	- 112 372	- 20.08	- 86.48	78.20%
		RF		92.9%	83.4%	- 323 544	- 3 937	- 13.65	- 28.28	98.80%
		XGB		93.4%	83.3%	- 383 004	-	- 15.14	-	100.00%
5	5164	CART	82.8%	92.3%	84.0%	- 1 966 473	- 46 323	- 76.83	- 22.31	97.60%
		RF		92.9%	84.0%	- 1 477 229	253 885	- 62.32	149.67	117.20%
		XGB		93.4%	84.1%	- 1 621 796	273 243	- 64.14	117.83	116.80%
6	6075	CART	29.1%	92.3%	89.7%	4 614 513	4 483 831	180.36	266.33	-2.80%
		RF		92.9%	89.9%	4 973 929	4 328 724	210.01	280.90	-13.00%
		XGB		93.4%	90.7%	5 354 770	5 368 917	211.69	301.57	0.30%
7	5534	CART	57.9%	92.3%	86.2%	2 312 231	3 310 314	90.36	412.71	43.20%
		RF		92.9%	86.1%	2 841 351	3 129 652	120.01	465.74	10.10%
		XGB		93.4%	86.6%	3 078 755	3 825 920	121.69	475.53	24.30%
8	5096	CART	84.5%	92.3%	83.8%	- 3 778 933	- 734 773	- 147.49	- 418.58	80.60%
		RF		92.9%	83.5%	- 2 513 261	8 914	- 106.03	20.13	100.40%
		XGB		93.4%	83.6%	- 2 920 405	34 492	- 115.47	45.75	101.20%
9	6452	CART	12.1%	92.3%	91.3%	52 090 240	47 706 070	2 034.15	2 170.64	-8.40%
		RF		92.9%	91.9%	46 171 160	42 049 900	1 949.05	2 082.36	-8.90%
		XGB		93.4%	92.5%	51 629 950	52 606 740	2 040.95	2 339.70	1.90%
10	4913	CART	64.9%	92.3%	85.2%	2 798 173	3 182 143	109.11	481.60	13.70%
		RF		92.9%	85.2%	2 502 903	2 743 070	105.69	554.76	9.60%
		XGB		93.4%	85.6%	2 920 720	3 438 303	115.40	576.64	17.70%
11	6199	CART	23.9%	92.3%	90.2%	9 335 438	8 527 444	364.60	454.78	-8.70%
		RF		92.9%	90.6%	8 570 307	7 931 029	361.80	460.42	-7.50%
		XGB		93.4%	91.2%	9 518 466	9 581 934	376.27	501.56	0.70%
12	4601	CART	89.5%	92.3%	83.6%	- 1 819 600	135 305	- 71.15	121.80	107.40%
		RF		92.9%	83.5%	- 1 575 489	159 620	- 66.48	215.65	110.10%
		XGB		93.4%	83.7%	- 1 668 346	228 226	- 65.99	208.69	113.70%
13	5379	CART	31.0%	92.3%	89.2%	4 160 423	3 882 623	162.44	241.06	-6.70%
		RF		92.9%	89.5%	4 018 432	3 666 219	169.65	249.54	-8.80%
		XGB		93.4%	90.0%	4 455 108	4 410 629	176.09	267.87	-1.00%
14	4887	CART	66.7%	92.3%	85.1%	1 858 140	2 575 538	72.44	442.86	38.60%
		RF		92.9%	85.0%	1 885 853	2 387 018	79.65	531.25	26.60%
		XGB		93.4%	85.4%	2 179 093	2 879 880	86.09	544.35	32.20%
15	6179	CART	24.4%	92.3%	90.3%	7 522 978	7 058 487	293.94	382.06	-6.20%
		RF		92.9%	90.6%	7 534 275	7 068 293	318.08	411.80	-6.20%
		XGB		93.4%	91.2%	8 219 857	8 265 167	324.94	442.88	0.60%
16	4627	CART	90.2%	92.3%	83.6%	- 2 506 749	- 139 983	- 97.95	- 121.44	94.40%
		RF		92.9%	83.5%	- 1 969 564	73 101	- 83.10	111.49	103.70%
		XGB		93.4%	83.6%	- 2 160 719	76 641	- 85.45	93.28	103.50%

CLV-based lapse management strategy

N°	time (s)	Model	% target diff	Accuracy		Retention gain		RG/target		Improvement
				$y_i$	$\tilde{y}_i$	$y_i$	$\tilde{y}_i$	$y_i$	$\tilde{y}_i$	
17	5679	CART	51.3%	92.3%	86.8%	5 220 695	5 811 833	203.94	583.55	11.3%
		RF		92.9%	86.9%	5 401 696	5 269 505	228.08	605.69	-2.40%
		XGB		93.4%	87.4%	5 943 841	6 682 230	234.94	670.03	12.40%
18	6074	CART	29.7%	92.3%	89.7%	3 862 156	3 397 247	150.95	203.11	-12.00%
		RF		92.9%	89.9%	4 127 224	3 550 730	174.26	230.67	-14.00%
		XGB		93.4%	90.6%	4 451 686	4 408 819	175.99	250.17	-1.00%
19	4528	CART	96.6%	92.3%	83.2%	- 2 929 448	- 85 465	- 114.46	- 1 482.06	97.10%
		RF		92.9%	83.3%	- 2 413 433	3 724	- 101.84	- 108.33	100.20%
		XGB		93.4%	83.3%	- 2 642 119	9 092	- 104.47	93.79	100.30%
20	5476	CART	60.9%	92.3%	85.9%	1 559 874	2 471 262	60.95	329.63	58.40%
		RF		92.9%	85.8%	1 994 645	2 517 111	84.26	422.45	26.20%
		XGB		93.4%	86.3%	2 175 670	3 089 897	85.99	422.77	42.00%
21	5906	CART	36.6%	92.3%	88.8%	705 721	922 490	27.69	60.21	30.70%
		RF		92.9%	88.9%	1 352 182	1 269 349	57.11	97.63	-6.10%
		XGB		93.4%	89.6%	1 342 882	1 428 722	53.09	96.76	6.40%
22	5278	CART	73.2%	92.3%	84.6%	- 1 596 562	169 852	- 62.31	41.78	110.60%
		RF		92.9%	84.6%	- 780 396	637 625	- 32.89	194.52	181.70%
		XGB		93.4%	85.0%	- 933 133	780 845	- 36.91	188.79	183.70%
23	4229	CART	99.2%	92.3%	83.3%	- 4 683 072	- 48 985	- 182.86	- 1 186.22	99.00%
		RF		92.9%	83.3%	- 3 536 046	-	- 149.21	-	100.00%
		XGB		93.4%	83.3%	- 4 000 747	-	- 158.17	-	100.00%
24	5391	CART	66.8%	92.3%	85.0%	- 3 062 732	- 388 289	- 119.35	- 80.44	87.30%
		RF		92.9%	84.7%	- 957 039	710 688	- 40.29	220.55	174.30%
		XGB		93.4%	85.3%	- 1 404 561	834 198	- 55.52	163.88	159.40%
25	5851	CART	43.0%	92.3%	87.8%	- 3 251 762	- 1 761 821	- 126.63	- 143.20	45.80%
		RF		92.9%	86.4%	- 1 013 089	79 273	- 42.69	11.90	107.80%
		XGB		93.4%	83.3%	- 1 582 006	4 396	- 62.52	287.68	100.30%
26	5182	CART	78.0%	92.3%	84.2%	- 5 554 044	- 1 491 522	- 216.63	- 549.23	73.10%
		RF		92.9%	83.6%	- 3 145 668	52 475	- 132.69	84.54	101.70%
		XGB		93.4%	83.3%	- 3 858 022	-	- 152.52	-	100.00%
27	5324	CART	32.7%	92.3%	88.9%	3 361 471	3 037 200	131.25	191.31	-9.60%
		RF		92.9%	89.3%	3 241 680	2 911 023	136.86	204.43	-10.20%
		XGB		93.4%	89.6%	3 596 593	3 546 671	142.15	222.04	-1.40%
28	4853	CART	70.3%	92.3%	84.7%	1 059 189	1 813 631	41.25	392.14	71.20%
		RF		92.9%	84.8%	1 109 101	1 808 616	46.86	474.33	63.10%
		XGB		93.4%	85.0%	1 320 578	2 141 271	52.15	482.34	62.10%
29	4746	CART	81.4%	92.3%	84.1%	- 1 165 404	458 223	- 45.60	172.83	139.30%
		RF		92.9%	84.0%	- 855 770	525 335	- 36.09	288.55	161.40%
		XGB		93.4%	84.1%	- 890 871	645 445	- 35.26	310.86	172.50%
30	4811	CART	73.9%	92.3%	84.3%	- 694 436	751 404	- 27.16	226.99	208.20%
		RF		92.9%	84.4%	- 13 512	1 018 369	- 0.51	356.48	7637.00%
		XGB		93.4%	84.7%	- 38 050	1 253 252	- 1.55	345.94	3393.70%
31	6243	CART	20.5%	92.3%	90.7%	14 747 500	13 838 380	576.23	690.22	-6.20%
		RF		92.9%	91.1%	14 816 830	13 378 460	625.54	743.34	-9.70%
		XGB		93.4%	91.5%	16 146 490	16 169 440	638.30	814.80	0.10%
32	5865	CART	41.4%	92.3%	88.0%	12 445 210	11 693 070	486.23	884.49	-6.00%
		RF		92.9%	88.3%	12 684 250	11 381 260	535.54	971.28	-10.30%
		XGB		93.4%	88.8%	13 870 470	14 101 470	548.30	1 048.38	1.70%

### 5.3. Comments

Several terms in the two previous tables need to be explained. "% target diff" represents how different  $y$  and  $\tilde{y}$  are. It is the percentage of subjects for which  $y_i = 1$  and  $\tilde{y}_i = 0$  : in other words, the proportion of lapseders not worth targeting with a given strategy. Then the table shows the 10-fold cross-validated mean accuracies, retention gains and RG/target with two methodologies: the columns denoted  $y_i$  represent the metrics obtained by a model with  $y_i$  as a response variable and accuracy as an evaluation metric, and the columns denoted  $\tilde{y}_i$  represent the metric obtained by a model with  $\tilde{y}_i$  as a response variable and  $RG$  as an evaluation metric.  $RG/target$  represents the achieved retention gain for every targeted individual, for  $y_i$ , it is  $RG / \sum_i \hat{y}_i$ , for  $\tilde{y}_i$  it is  $RG / \sum_i \hat{\tilde{y}}_i$ . Eventually, "Improvement" represents the percentage of improvement between the  $RG$  obtained with a classification on  $y_i$  and the gain obtained with a classification on  $\tilde{y}_i$ .<sup>1</sup>

Some LMS are worth focusing on. For every strategy, we can display its results when trained on random samples of the dataset: the boxplots below summarize some typical key results illustrated by several strategies. Those results will be discussed in the Section 6.

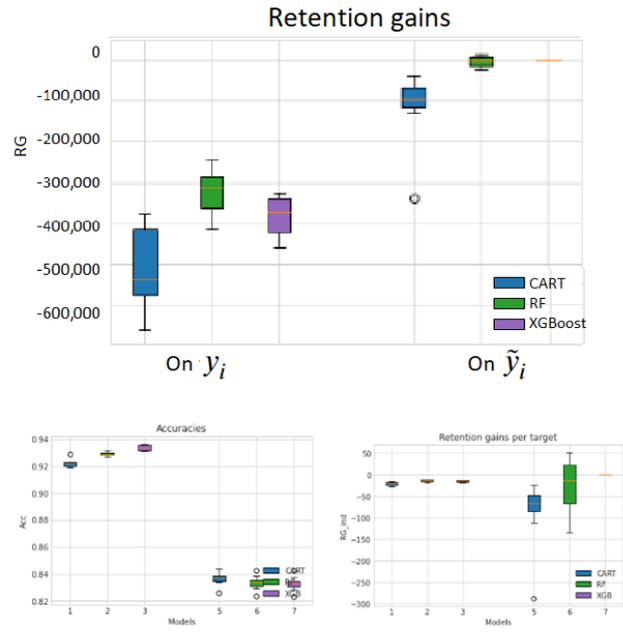


Figure 3: Strategy n°4: (Very negative result on  $y_i$  and a loss-limiting result on  $\tilde{y}_i$ .)

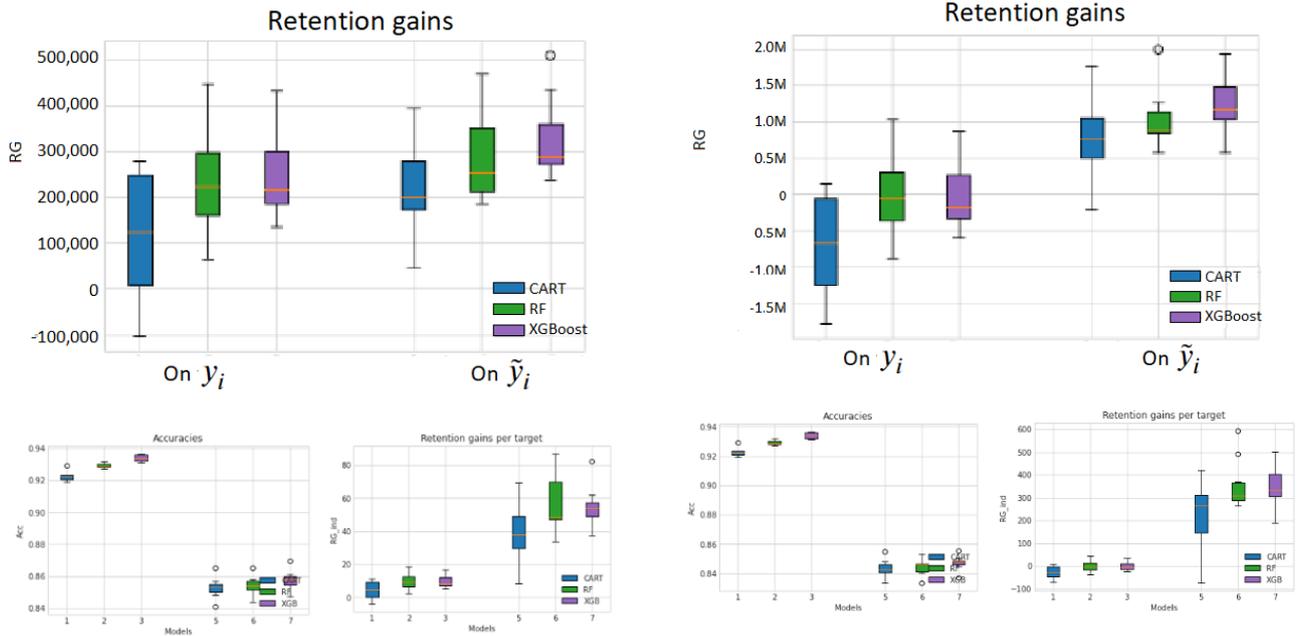


Figure 2: Strategy n°1: (Positive result on  $y_i$  and an improved result on  $\tilde{y}_i$ .)

Figure 4: Strategy n°30: (Negative result on  $y_i$  and positive one on  $\tilde{y}_i$ )

<sup>1</sup>In order to account for negative retention gains, the improvement is computed with an absolute value for the denominator. This leads to a rather unintuitive improvement measure whenever one of the models yields negative RG and the other positive RG.

**Remark 6.** With considerable computation power and great parallelization, the results for all strategies - see other strategies in Appendix A.7 - were obtained with a wall time of less than 4 days and a CPU time of more than 100 days.

## 6. Discussion

### 6.1. General statements

RF and XGBoost perform globally better than CART in terms of mean accuracy and RG. It is true for all 32 LMS considered in Table 1. Globally, XGBoost is more consistent and is the best model in most scenarios, both with and without the CLV-based measure. It is only outperformed by RF in strategies n°25 and 26.

The vast majority of strategies, including all the realistic ones, show that a classification on  $\tilde{y}_i$  produces a targeting that yields better RG than a classification on  $y_i$ . Conversely, a classification on  $y_i$  produces a targeting that yields better accuracies regarding whether a policyholder will churn than a classification on  $\tilde{y}_i$ . These results were expected because of the models' respective objectives. Even if it is not surprising, it once again shows that for an insurer, lapse prediction and lapse management strategy are two very different prediction problems, often treated as similar ones.

Our CLV-augmented modelization shows different behaviour depending on the strategy considered. As highlighted by Figure 2, a model on  $y_i$  is greatly improved by our framework regarding RG and RG/target. Conversely, its accuracy in lapse prediction is not optimal.

An attractive property of our framework can be observed in Figure 3: it yields loss-limiting targeting. When the LMS considered is too aggressive, it will usually prefer to predict that an LMS should not be applied at all ( $\forall i, \hat{y}_i = 0$ ), thus generating a *RG* around 0€. This is made evident in some extreme strategies (like LMS n°23 for instance).

On less extreme strategies, it shows to yield substantial improvement when classification on  $y_i$  gives negative *RG*. That observation confirms what was already pointed out by Loisel et al. (2021): it can even turn a negative *RG* into a positive one (see LMS n°5, 12, 22, 24, 29 and 30 - Figure 4)

Our framework also improves a strategy where a classification on  $y_i$  gives high *RG*. However, the improvement decreases as the difference between the number of lapsers and the number of lapsers that would be profitable if retained is sizeable. An example of that is shown in Figure 5.

Finally, we can generate LMS for which our framework does not improve the expected *RG*. It is the case in LMS n°13, 18 or 27 (See e.g Figure 6). In LMS n°13, we can see that the mean of the *RG* is not improved, but the median is. In all those cases, the *RG* per target produced by the CLV-augmented model is greatly improved, indicating that a CLV-augmented strategy prefers to target fewer policyholders but only those who would generate high future profits.

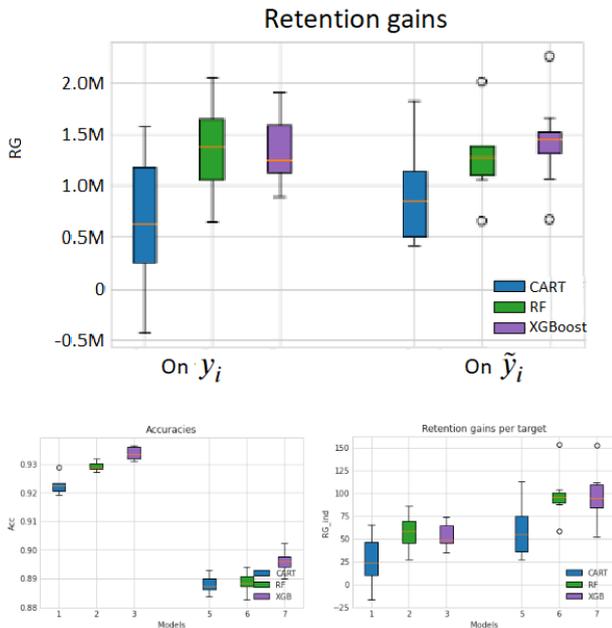


Figure 5: Strategy n°21: (High positive result on  $y_i$  slightly improved with  $\tilde{y}_i$ .)

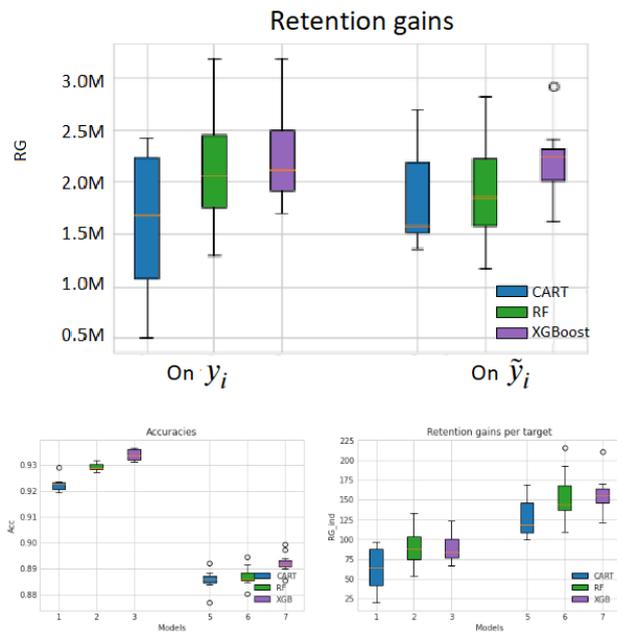


Figure 6: Strategy n°13: (Results on  $y_i$  better than results on  $\tilde{y}_i$ .)

This last observation explains why a CLV-augmented LMS generates higher RGs when the cost of contact  $c$  is considerable. Indeed, the more costly a contact is, the more precise and specific a targeting strategy should be.

Generally, if we consider our various LMS, excluding LMS n°30 that has a very high improvement ratio:

The average observed RG improvement of a CLV-augmented framework over the classical lapse one is 57,9%<sup>a</sup>. If we weight these results by the expected RGs, the average improvement is still 31,7%.

<sup>a</sup>Using XGBoost

### 6.2. Marketing decision making

We already pointed out that the improvement of a lapse management strategy including CLV grows with the proportion of lapsed with a negative CLV (see Figure 26 in Appendix A.7). Models resulting from our framework do not consider them as good targets. In fact, there is a Pearson correlation coefficient of 77% between RG improvement and the proportion of target differences among the LMS detailed in Table 1<sup>2</sup>. Of course, as the improvement ratio has no clear interpretation in some cases, this analysis should be carried out in more depth, separating the cases where both RG - with and without the inclusion of CLV - are positive from the cases where one of them is negative. By doing so, we observe that the Pearson correlation coefficient for LMS yielding positive RG regardless of the inclusion of CLV is even higher: 83%.

In terms of targeting, it seems crucial to understand what differentiates a subject for which  $y_i = 1$  and  $\tilde{y}_i = 0$  from the others. An investigation of such policyholder's profile can be carried out for every lapse management strategy. We take the example of LMS n°1, where 62,6% of policyholders were in that case (see Section 5.2). With that strategy, the profile of non-targeted lapsed indicates that :

- 57.2% of them are men, similar to the entire dataset,
- 76.4% of them contracted product n°1 whereas 72% of all policyholders chose it,
- the mean seniority of their policy is 10.4 years compared to the 13.4 years for the complete dataset,
- the mean face amount of such policies is 12,156, whereas the average face amount for all considered policies is 40,263.

In that strategy, our framework indicates that marketing efforts on low seniority policyholders with low face amount policies are inefficient. Of course, this conclusion is only valid for the considered LMS; however, our framework allows us to conduct such analysis for any LMS and interpret the results at an individualized level.

<sup>2</sup>Taking the results of XGBoost and excluding LMS n°30 that has a very high improvement ratio.

### 6.3. Management rules decision making

Sensitivity analysis of those results can highly benefit management rules decision-making. This framework serves as a tool that compares future hypothetical lapse management strategies in order to choose the best one - among realistic scenarios -. It can also be used to tune a given strategy by answering questions like:

- For which incentive  $\delta$  the retention strategy becomes profitable ?
- For which acceptance probability  $\gamma$  the retention strategy becomes profitable ?
- With a given budget, what is the optimal list of policies that should be targeted?
- At which horizon  $T$ , the retention strategy become profitable ? In other words, when can the insurer expect a return on investment?

Answering these questions constitutes a 1-parameter sensitivity analysis. In our framework, six parameters influence the expected retention gain ( $p, \delta, \gamma, c, d, T$ ).

We can argue that among them are three structural parameters that are insurer's dependent and not linked to the external state of the world:  $\delta, \gamma$  and  $c$ . Among them, the contact cost  $c$  is more or less fixed and can not be easily changed by the insurer. Conversely,  $\delta$  and  $\gamma$  are to be chosen by the insurer. Moreover, they also are correlated with management and commercial efficiency - an efficient campaign impacts the final  $\gamma$  - and correlated together: the higher the incentive  $\delta$ , the higher the probability of acceptance  $\gamma$ .

By fixing all other parameters and trying various combinations of  $\delta$  and  $\gamma$  we obtain the following 3D surfaces:

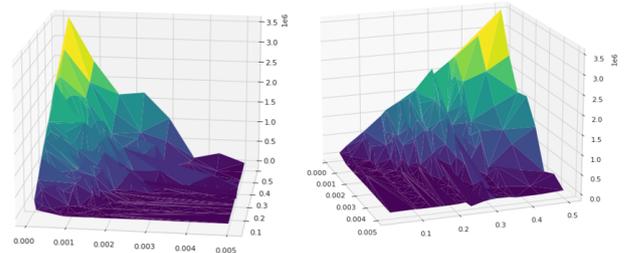


Figure 7: 3d plot ( $\delta, \gamma, RG$ )

This surface is not surprising and indicates that the higher the acceptance rate and the lower the incentive, the higher the retention gain. The surface gradient can give powerful insights on the most efficient commercial efforts to make: is it better for the insurer to propose lower incentives and manage to conserve the same acceptance probability or to put commercial effort into improving the acceptance probability for the same proposed incentive? This surface directly addresses this question.

**Remark 7.** Of course, the interdependency of those parameters should make some part of this surface unrealistic from a management decision-making point of view. The insurer

should consider such dependencies when designing a lapse management strategy.

Among the six parameters are also three conjectural parameters that depend on the external state of the world: the insurer's profitability  $p$  (that depends on competition, macroeconomic considerations or regulation), the discount rate  $d$  and the time horizon  $T$  (that can be driven by the insurer's vision but also by regulation<sup>3</sup>). Among them, we chose to fix  $p$  and let  $d$  and  $T$  vary. Moreover,  $p$  and  $T$  are obviously interdependent and considered through the management's prospective view of the conjecture's evolution. A given interest rate scenario should represent a curve on the following surface:

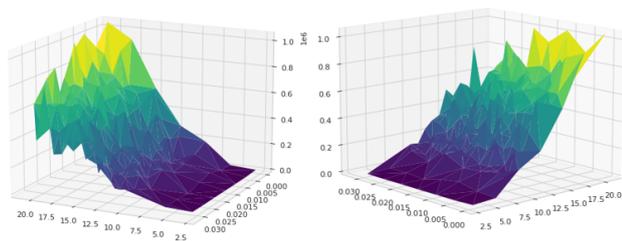


Figure 8: 3d plot ( $d$ ,  $T$ ,  $RG$ )

This surface is less smooth than the one displayed in Figure 7 and seems to indicate a more unstable relationship between  $RG$  and the conjectural parameters. An explanation of that behaviour can be that those surface points are generated by running our framework on a random subsample of our dataset, for computation time considerations. Generating the same surface with more policyholders is likely to give a smoother behaviour.

**Remark 8.** Of course, the interdependency of  $T$  and  $d$  should make some part of this surface unrealistic from an actuarial point of view. Actuarial rate projections would give precise plausible scenarios on this surface. Such considerations should be taken into account by the insurer when designing a lapse management strategy.

**Remark 9.** The insurer can also use our framework to measure the retention gain to be expected at different time horizons obtained by existing retention campaigns. In that case, the insurer would have to neutralize the effect of the existing LMS in order to estimate the control portfolio's future value. We leave this remark as future work for applied risk management research.

## 7. Conclusion, limitations and future work

The work carried out in this paper shows that including CLV in lapse management strategy can largely benefit an insurer's decision-making ability regarding lapse management strategy. We showed that survival tree-based models can outperform parametric approaches in such actuarial contexts. Then, our comparison of tree-based models on different lapse management strategies indicated that our CLV-based framework leads to increased predicted gains for any realistic scenario and acts as a loss-limiting targeting approach, regardless of the retention strategy. Finally, the discussion section highlighted the fact that our modelization can give insights to the life insurer regarding commercial and strategic decision-making.

The framework and methodologies described in this paper suffer some limitations. For instance, we can argue that following one single fixed strategy for every policyholder is unrealistic. We could easily imagine an extension of our models to individualized lapse management strategies that would vary between subjects, but that could also be adjusted with time. There is also room for improvement regarding the correlations of LMS parameters: gamma and delta are evidently interdependent parameters for an insurance company, and this interdependency could be considered.

This paper defines a practical management tool for life insurers as those models can be used to measure the  $RG$  and improve real strategies used in existing retention campaigns. Finally, our vision of CLV, and by extension, our whole methodology design could be improved by using longitudinal data that would yield time-dynamic results. We leave those two last observations for future work.

A real-life comparison between an actual retention strategy targeting and both the naïve and CLV-improved methodologies could be insightful for the insurer<sup>4</sup>.

## Acknowledgments

Work(s) conducted within the Research Chair DIALog under the aegis of the Risk Foundation, an initiative by CNP Assurances. The authors would like to express their very great gratitude to Marie Hyvernaud and Stéphanie Dosseh for their valuable and constructive suggestions during the development of this research work. Special thanks should be given to Marie Hyvernaud for her contribution to code writing.

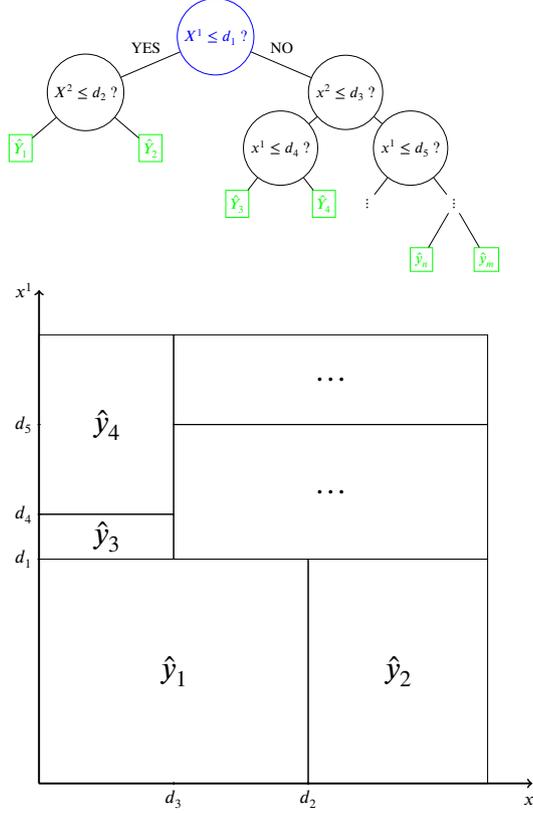
<sup>3</sup>The ORSA time horizon with the strategic and the long-term business planning time horizon should be both considered.

<sup>4</sup>On that point, it is to be noted that this use of our framework is not trivial. As  $RG$  is computed as the difference between a control portfolio value and a lapse-managed one, this setting would first require estimating a  $CPV$  that was never observed.

## A. Appendix

### A.1. CART

For visualization's sake, we can show below an example of a CART model partitioning  $(x^1, x^2)$ , the space defined by two covariates along different thresholds  $(d_1, d_2, \dots)$ :



Generally speaking, there are far more than two covariates which makes the space partitioning representation difficult. Summarizing this cutting with a binary tree is obviously a very interpretable way of displaying the modelization: a node  $g$  is identified with a subregion of  $\mathcal{X}$  it represents. The first step of fitting a CART tree on a dataset  $\mathcal{D} = (\mathbf{x}_i, y_i)_{i=1}^n$ , where  $\mathbf{x}_i$  is the vector of covariates for the  $i$ -th subject and  $y_i$  its response variable is applying a splitting procedure until a stopping criteria is met. Various splitting criteria exist depending on the nature of the prediction problem to solve. As we will work with both classification and regression settings, it seems natural to explicit both splitting mechanisms here.

In a classification context, let's define  $p_i, i \in \{1, \dots, P\}$  as the proportion of observation of class  $i$  in  $\mathcal{D}$ . We extend this idea by defining  $p_i(g)$  as the proportion of observation of class  $i$  in  $\mathcal{D}(g)$ , the collection of observations in node  $g$ . An impurity function  $\phi$ , is a function defined for  $p_i, i \in \{1, \dots, P\}$ , with  $p_i \geq 0$  and  $\sum_i p_i = 1$  such that  $\phi(p_1, \dots, p_P)$  is positive, the minimum of  $\phi$  is reached whenever any of the  $p_i = 1$ , then  $\phi(p_1, \dots, p_P) = 0$ , its maximum is reached for  $\phi(\frac{1}{P}, \dots, \frac{1}{P})$  and it is symmetric with regard to its arguments.

For CART, usual classification impurities are:

- Entropy :  $\phi(p_1, \dots, p_P) = - \sum_i p_i \log(p_i)$
- Gini index:  $\phi(p_1, \dots, p_P) = 1 - \sum_{i < j} p_i p_j$

For a given node  $g$  with  $\mathcal{N}(g)$  observations, the impurity of the node is given by  $I(g) = \phi(p_1(g), \dots, p_P(g))$ .

At each node of a CART, the optimal split is chosen as the split that reduces the impurity the most. That is to say, the split that maximizes the following gain function by splitting the parent node  $g_p$  into the two child nodes  $g_l$  and  $g_r$ :

$$G(g_p; g_l, g_r) = I(g_p) - \left( \frac{\mathcal{N}(g_l)}{\mathcal{N}(g_p)} I(g_l) + \frac{\mathcal{N}(g_r)}{\mathcal{N}(g_p)} I(g_r) \right),$$

This splitting procedure allows growing the tree until a splitting criterion - such as a maximal depth or a minimal number of observations in the terminal nodes - is met. A tree grown following this step is called a maximal tree.

Usually, the maximal tree is not an efficient model as it leads to a very thin partitioning of  $\mathcal{X}$  that does not produce good predictions for test observations. The bias of the maximal tree is low, but its variance is very high. Breiman suggested a method to select an optimal subtree of the maximum tree that would be deep enough to produce low-bias prediction but small enough to generalize well to new data. The idea is to penalize the performance of the subtrees by the number of terminal nodes they have, allowing to construct a sequence of optimal subtrees depending on the strength of the penalization. That pruning phase intends to find the sequence of best trees of  $1, 2, \dots, l$  leaves. The optimal tree is then selected among this sequence, by cross-validation.

This method leads to a clear and interpretable modelization that is not sensitive to outliers and can handle missing values. This makes CART a valid choice of modelization for real business applications. In contrast, the predictions obtained with CART are very sensitive to changes in the training set, which makes it an unstable model. This lack of stability led to the design of more elaborate procedures to design tree-based models, such as the *Bagging* and *Boosting* techniques detailed in the following sections.

### A.2. Random forest

In the random forest algorithm,  $q$  covariates are randomly selected among the  $p$  available features at each node. At each node, only those  $q$  covariates are considered to find the optimal split.

Let's define  $\Theta_{obs}$ , the random variable representing the bootstrap over observations of  $\mathcal{L}$  and  $\Theta_{cov}(q)$ , the random variable representing the random drawings of  $q$  covariates over the  $p$  available features at each node. We assume that  $\Theta_{obs}$  and  $\Theta_{cov}(q)$  are independent of  $\mathcal{L}$ . The  $i$ -th realisations

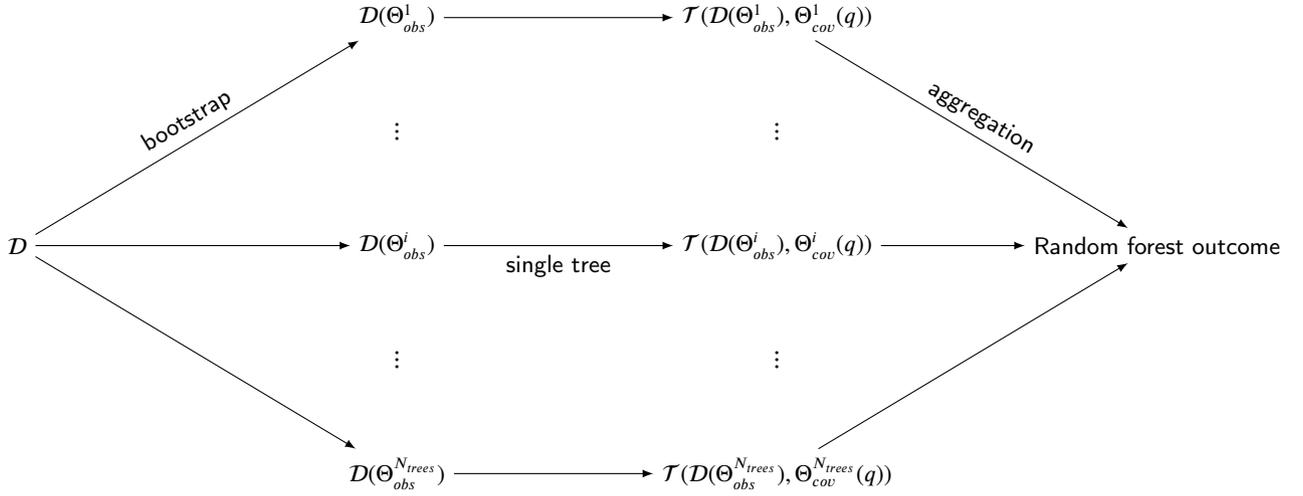


Figure 9: Random forest mechanism

of those bootstrap steps are denoted  $\Theta^i_{obs}$  et  $\Theta^i_{cov}(q)$ .

$\mathcal{L}(\Theta^i_{obs})$  would then be a subsample of  $\mathcal{L}$  and  $\mathcal{T}(\mathcal{L}(\Theta^i_{obs}), \Theta^i_{cov}(q))$  the maximal tree grown with CART algorithm from this subsample with the  $p$  features considered at each node, selected with  $\Theta^i_{cov}(q)$ . The mechanism of a random forest built over  $N_{trees}$  trees is summarized below:

The parameter  $q$  is fixed in advance and is constant among trees of the same forest. The fact that the random feature selection happens at each node of a tree ensures that every tree deep enough is grown by considering almost surely every covariate. This algorithm, with two different bootstrap procedures, has been developed by Breiman (1996) and Amit and Geman (1997) and synthesized in Breiman (2001).

When predicting the forest outcome for an observation, the aggregation step consists of propagating the observation through every tree in the forest. The prediction of the forest is the majority vote among all trees of the mean outcomes of each individual tree respectively, in a classification or a regression context. In other words, let  $h_i(x) = h_i(x, D(\Theta^i_{obs}), \Theta^i_{cov}(q))$  denote the prediction of the observation  $x$  given by the  $i$ -th tree, grown with the bootstrap realizations  $\Theta^i_{obs}$  and  $\Theta^i_{cov}(q)$ , the prediction of the random forest is  $RF(x)$ :

$$RF(x) = \text{mode}(h_1(x), \dots, h_{N_{trees}}(x)), \text{ for classification}$$

$$RF(x) = \frac{1}{N_{trees}} \sum_{i=1}^{N_{trees}} h_i(x), \text{ for regression}$$

Theoretical guarantees on random forests exist and often rely on the hypothesis that all the trees within the forest are independent. The bootstrap procedure aims at diminishing the correlations between different trees to get closer to that

assumption.

### A.3. XGBoost

If tree-based boosting approaches share with bagging the global concept of aggregating weak predictions, the construction of each tree is very different. For boosting techniques, it relies on a recursive relation: a first tree is grown, then a second one that will put more weight on poorly predicted observations, and so on: single trees are grown iteratively such that each subsequent tree aims at reducing the errors of the previous one. Intuitively, the goal of boosting thought iterations is to focus on regions of the original dataset that yield poor predictions.

All those weak single trees build iteratively and then aggregated with a weighted mean, with weights related to their accuracy.

Generally, let  $\mathbf{w}^i = (w_1, \dots, w_n), \forall i \in [1, \dots, N_{trees}]$  be the vector of observations weights during the  $i$ -th iteration of the boosting algorithm. We initialize:  $\mathbf{w}^1 = (\frac{1}{n}, \dots, \frac{1}{n})$ , let  $D^{\mathbf{w}^i}$  be the original dataset, with observations weighted by  $\mathbf{w}^i$ . Let  $\mathcal{T}(\mathbf{w}^i)$ <sup>5</sup> be the tree grown from  $D^{\mathbf{w}^i}$ . The final model is denoted  $B(D)$ .

"The predictions are combined in an additive manner, where the addition of each base model improves (or "boosts") the overall model. Therefore, the overall model  $f$  is an additive model of the form:"

$$f(\mathbf{x}) = \sum_{m=1}^M \beta_m g(\mathbf{x}; \theta_m),$$

"where  $M > 0$  denotes the number of base learners, and  $\beta_m \in \mathcal{R}$  is a weighting term. The function  $g$  refers to a base learner and is parameterized by the vector  $\theta$ . Individual

<sup>5</sup>This notation is clearly simplified and does not take other parameters into account

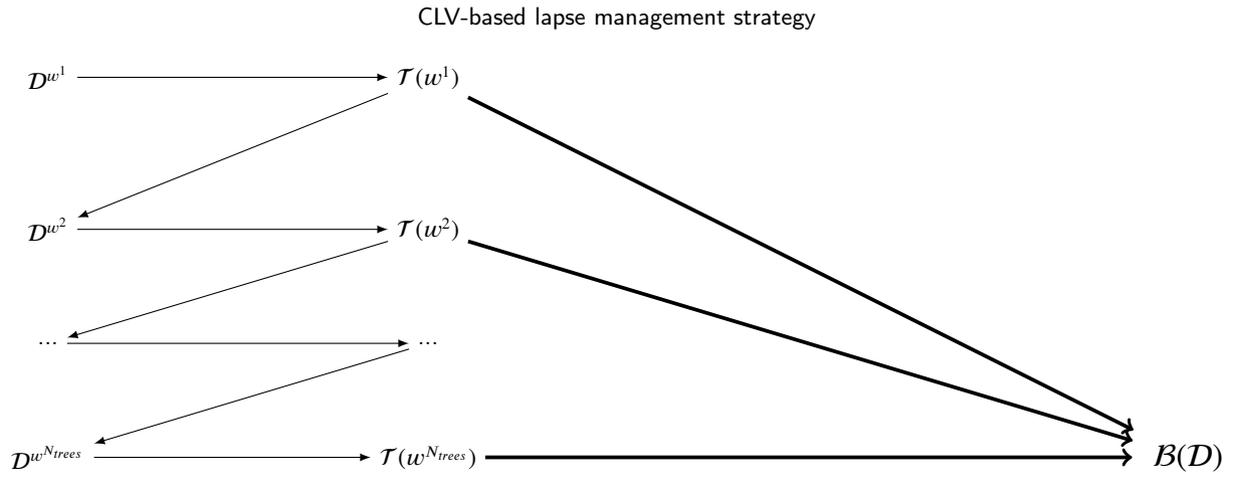


Figure 10: Tree-boosting mechanism

base learners differ in the configuration of their parameters  $\theta$ , which is indicated by a subscript  $m$ ."

We can illustrate the mechanism of a boosting algorithm like so:

There are many boosting techniques adapted to trees, and the main difference between them can be their method of weighting training data points or the hypotheses they rely on. In this paper, we will only consider the library *XGBoost* as it is regarded as one of the most competitive boosting algorithms in terms of prediction accuracy, computation time, and flexibility *XGBoost* has been developed by Chen and GuestrinChen and Guestrin (2016) and relies on a gradient boosting procedure. One of the main advantages of *XGBoost* resides in its optimization, as it can be finely tuned through numerous parameters and benefits significantly from parallel computation.

#### A.4. Confusion matrix and evaluation metrics

Obviously, with a binary outcome, an observation can be well classified or misclassified in only four ways: a churned observation can be classified as such or misclassified, and an active observation can be classified as such or misclassified. A confusion matrix is essentially a table containing the number of observations  $N(j, k)$  in the test set that fall in each of those possibilities:

$$N(j, k) = \sum_{i=1}^N \mathbf{1}(y_i = j, \hat{y}_i = k).$$

		Real outcome $y$		Total
		1	0	
Predicted outcome $\hat{y}$	1	$N(1, 1)$	$N(0, 1)$	$N(-, 1)$
	0	$N(1, 0)$	$N(0, 0)$	$N(-, 0)$
Total		$N(1, -)$	$N(-, 0)$	$n$

$N(1, 1)$ ,  $N(1, 0)$ ,  $N(0, 1)$ ,  $N(0, 0)$  are respectively called the number of true positive, false negative, false positive

and true negative.

From this number, we can derive several well-used metrics:

$$\begin{aligned} \text{Accuracy}(y, \hat{y}) &= \frac{N(1, 1) + N(0, 0)}{n} \\ \text{Precision}(y, \hat{y}) &= \frac{N(1, 1)}{N(1, 1) + N(0, 1)} = \frac{N(1, 1)}{N(-, 1)} \\ \text{Recall}(y, \hat{y}) &= \frac{N(1, 1)}{N(1, 1) + N(1, 0)} = \frac{N(1, 1)}{N(1, -)} \\ \text{F}_1\text{Score}(y, \hat{y}) &= \frac{N(1, 1)}{N(1, 1) + \frac{N(1, 0) + N(0, 1)}{2}} = \frac{2}{\frac{1}{\text{recall}} + \frac{1}{\text{precision}}} \\ &= 2 \cdot \frac{\text{recall} \cdot \text{precision}}{\text{recall} + \text{precision}} \end{aligned}$$

$\text{Accuracy}(y, \hat{y})$  is undoubtedly the most intuitive performance measure, and it is defined as the proportion of correctly predicted observations over all observations. It is widely used for churn analysis but appears to be a satisfying performance measure for balanced datasets, which is debatable in our case.

$\text{Precision}(y, \hat{y})$  measures the proportion of observations that actually churned among the observations predicted as churners: the higher this metric, the lower the false positive rate.

$\text{Recall}(y, \hat{y})$  measures the proportion of observations that were correctly identified as churners among the actually churned observations. The higher this metric, the lower the false-negative rate, which is of great interest for churn predictions that lead to the application of a retention strategy. The real-world retention strategies are often limited by budget or deployment constraints; a low recall would mean that a significant part of the retention budget was spent on policyholders that were not churners. It is often preferable to target a smaller number of policyholders with a more substantial proportion of churners at the risk of missing a lot of them than targeting a more significant number of policyholders with a lower proportion of churners, even if that means that more churners were identified as such.

Lastly,  $F_1$  Score can be seen as a weighted average between precision and recall. It measures a model's ability to predict churners in terms of both precision and recall. It is usually more efficient than accuracy in the presence of uneven class distribution. Accuracy seems to perform better if false positives and false negatives have similar costs, which is arguably not the case in our study.

### A.5. Competing risk framework

There are several regression models to estimate the global hazard and the hazard of one risk in settings where competing risks are present: modeling the cause-specific hazard and the subdistribution hazard function. They account for competing risks differently, obtaining different hazard functions and thus distinct advantages, drawbacks, and interpretations. We will here quickly introduce those approaches' theoretical and practical implications and justify which one we will use in our modelizations.

In Cause-specific regression, each cause-specific hazard is estimated separately, in our case, the cause-specific hazards of lapse and death, by considering all subjects that experienced the competing event as censored: We remind that  $J_T = 0$  corresponds to an active subject that did not experience lapse  $J_T = 1$  or death  $J_T = 2$ . The cause-specific hazard rates regarding the  $j$ -th risk ( $j \in [1, \dots, J]$ ) are defined as:

$$\lambda_{T,j}(t) = \lim_{dt \rightarrow 0} \frac{P(t \leq T < t + dt, J_T = j \mid T \geq t)}{dt}$$

We can recover the global hazard rate as  $\lambda_{T,1}(t) + \dots + \lambda_{T,J}(t) = \lambda_T(t)$ , and derive the global survival distribution of  $T$  as

$$\begin{aligned} P(T > t) &= 1 - F_T(t) = S_T(t) \\ &= \exp\left(-\int_0^t (\lambda_{T,1}(s) + \dots + \lambda_{T,J}(s)) ds\right) \end{aligned}$$

This approach aims at analysing the cause-specific "distribution" function:  $F_{T,j}(t) = P(T \leq t, J_T = j)$ . In practice, it is called the Cumulative Incidence Function (*CIF*) for cause  $j$  and not a distribution function since  $F_{T,j}(t) \rightarrow P(J_T = j) \neq 1$  as  $t \rightarrow +\infty$ . By analogy with the classical survival framework, the *CIF* can be characterized as  $F_{T,j}(t) = \int_0^t f_{T,j}(s) ds$ <sup>7</sup>, where  $f_{T,j}$  is the improper<sup>8</sup> density function for cause  $j$ . It follows that:

$$f_{T,j}(s) = \lim_{dt \rightarrow 0} \frac{P(t \leq T < t + dt, J_T = j)}{dt} = \lambda_{T,j}(t) S_T(t)$$

The equation above is self-explanatory: the probability of experiencing cause  $j$  at time  $t$  is simply the product of

<sup>7</sup>We suppose that  $T$  has a continuous distribution

<sup>8</sup>Because derived from the *CIF*, an improper cumulative distribution function

surviving the previous time periods by the cause-specific hazard at time  $t$ . We finally obtain the *CIF* for cause  $j$  as:

$$F_{T,j}(t) = \int_0^t \lambda_{T,j}(s) \exp\left(-\int_0^s \lambda_T(u) du\right) ds.$$

There are several advantages to that approach. First of all, cause-specific hazard models can be easily fit with any classical implementation of CPH by simply considering as censored any subject that experienced the competing event. Then the *CIF* is clearly interpretable and summable  $P(T \leq t) = F_{T,1}(s) + \dots + F_{T,J}(s)$ <sup>9</sup>. On the other hand, the *CIF* estimation of one given cause depends on all other causes: it implies that the study of a specific cause requires estimating the global hazard rate, and interpreting the effects of covariates on this cause is difficult. Indeed, part of the effects on a specific cause comes from the competing causes, but in our setting, we are only interested in the prediction of the survival probabilities, not their interpretation as such.

We have introduced it at the beginning of this section; another approach is often considered to analyze competing risks and derive a cause-specific *CIF*. This other approach, called the subdistribution hazard function of Fine and Gray regression, works by considering a new competing risk process  $\tau$ . Without loss of generality, let's consider death as our cause of interest:

$$\tau = T \times \mathbb{1}_{J_T=2} + \infty \times \mathbb{1}_{J_T \neq 2}.$$

It has the same as  $T$  regarding the risk of death,  $P(\tau \leq t) = F_{T,2}(t)$  and a mass point at infinity  $1 - F_{T,2}(\infty)$ , probability to observe other causes ( $J_T \neq 2$ ) or not to observe any failure. In other words, if the previous approach considered every subject that experienced competing events as censored, this approach considers a new and artificial at-risk population. This last consideration is made clear when deriving the hazard rate of  $\tau$ :

$$\lambda_\tau(t) = \lim_{dt \rightarrow 0} \frac{P(t \leq T < t + dt, J_T = 2 \mid \{T \geq t\} \cup \{T \leq t, J_T \neq 2\})}{dt}$$

Finally, we obtain the *CIF* for the risk of death as:

$$F_{T,2}(t) = 1 - \exp\left(-\int_0^t \lambda_\tau(s) ds\right).$$

This subdistribution approach resolves the most important drawback to cause-specific regression, as the coefficients resulting from it do have a direct relationship with the cumulative incidence: estimating the *CIF* for a specific cause does not depend on the other causes, which makes the interpretation of *CIF* easier. The subdistribution hazard models can be fit in R by using the *crr* function in the *cmprsk* package or using the *timereg* package. Still, to our knowledge, there is no implementation of a Fine and Gray model in Lifelines or, more generally, Python. We can also

<sup>9</sup>unlike to the function  $1 - \exp\left(-\int_0^t \lambda_{T,j}(u) du\right)$ .

note that these two approaches are linked, Putter et al. and the link between  $\lambda_\tau(t)$  and  $\lambda_{T,j}(t)$  is given by:

$$\lambda_\tau(t) = r_j(t)\lambda_{T,j}(t), \text{ with } r_j(t) = \frac{P(J_T = 0)}{\sum_{p \neq j}^J P(J_T = p)}$$

In other words, if the probability of any competing risk is low, the two approaches give very close results.

### A.6. Survival modelizations results

$r_{i,t}^{lapse}$  represents the probability that the policy of subject  $i$  is still active at time  $t$ , given that it was active at its last observed time. Predicting the overall conditional survival with the competing risks, in that case, can be achieved by creating a combined outcome. The policy ends with death or lapse, whichever comes first, and to compute  $r^{lapse}$ , we recode the competing events as a combined event. In terms of statistical guarantees, this approach is compatible with any survival analysis method.

In the following sections of this appendix,  $r_{i,t}^{acceptant}$  indicates the probability of survival for subject  $i$  at time  $t$  given that it will not lapse. In other words, it is the survival probability regarding only the risk of death. As detailed in Section 4.1.1, this corresponds to the cause-specific survival probability for death. It is to be noted that the density from which we derive our survival probabilities is improper as it derives itself from the *CIF*, which is not a proper distribution function.<sup>10</sup> Therefore, any conclusion about those probabilities should be drawn with care. Our modelization is derived from a Cox cause-specific hazard regression with a spline baseline hazard from the Python library Lifelines. Similarly to  $r^{lapse}$ , covariates selection and tuning are performed by minimizing AIC.

#### A.6.1. Cox-model

We first decide to estimate survival with a Cox Proportional hazard model with a spline baseline hazard from the Python library Lifelines. Covariate selection and tuning are performed by minimizing AIC. Here is what  $r^{acceptant}$ , the vector of cause-specific probabilities, looks like, and we can compare it to  $r^{lapse}$  on some subjects:

<sup>10</sup>as it does not tend to 1 as  $t$  goes to  $+\infty$

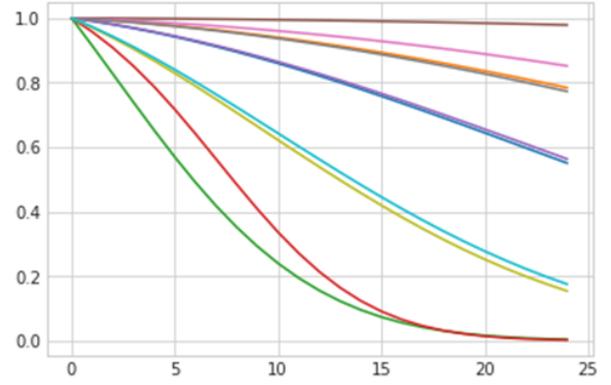


Figure 11: 10 policyholders' survival curve for  $r^{acceptant}$  with Cox model

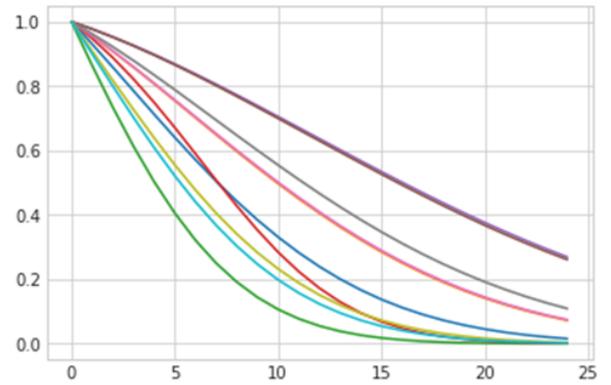


Figure 12: 10 policyholders' survival curve for  $r^{lapse}$

The effect of various covariates on the survival outcome can be found below:

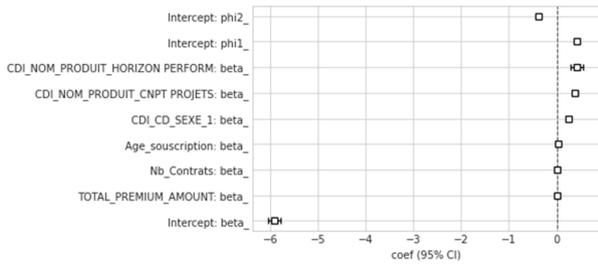


Figure 13: Coefficient plot for  $r^{lapse}$

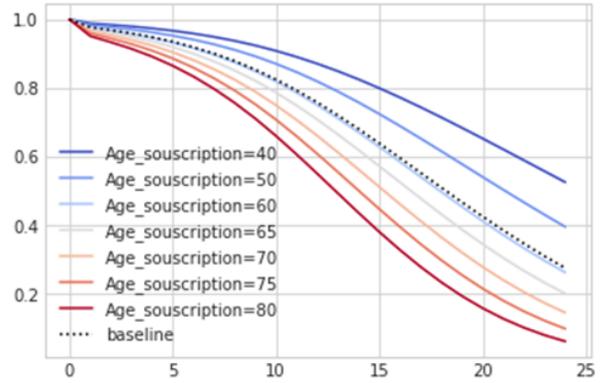


Figure 16:  $r^{lapse}$  trajectories for different ages

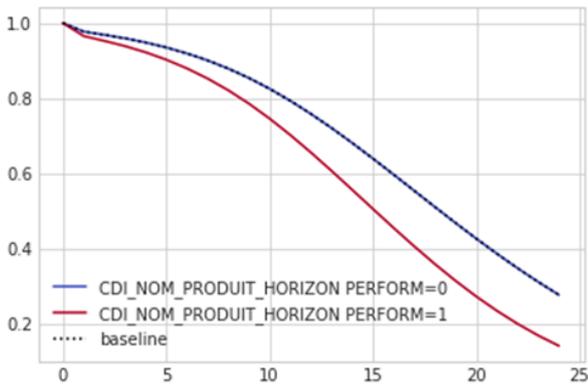


Figure 14:  $r^{lapse}$  trajectories for different products

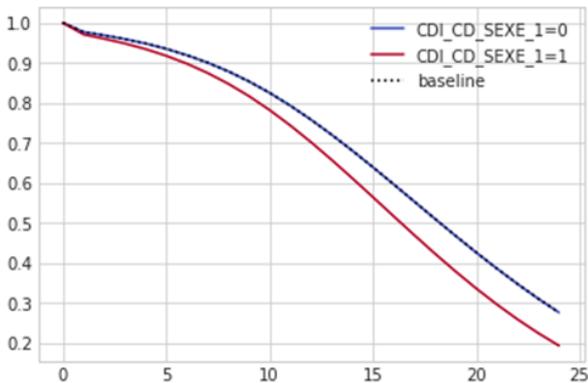


Figure 15:  $r^{lapse}$  trajectories by gender

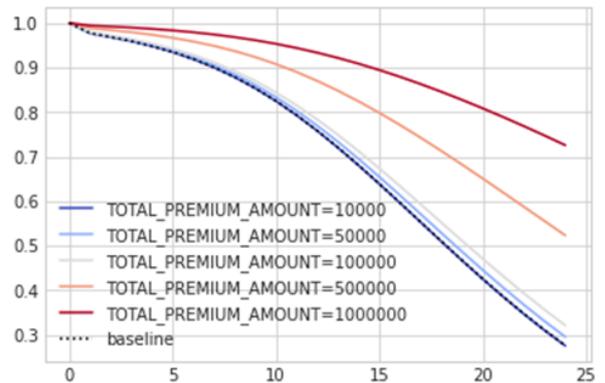


Figure 17:  $r^{lapse}$  trajectories for different face amounts

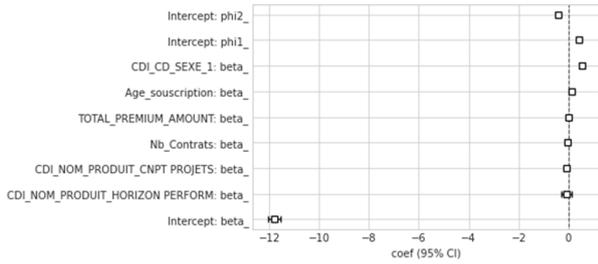


Figure 18: Coefficient plot for  $r^{acceptant}$

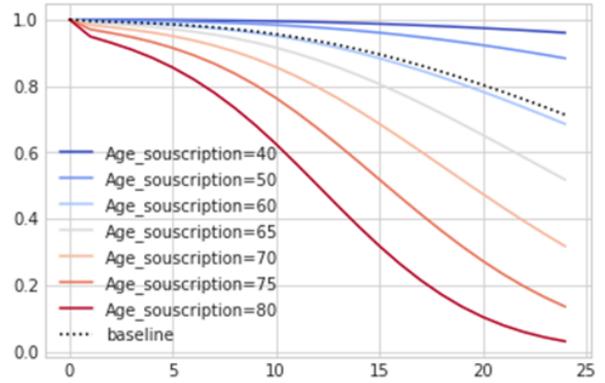


Figure 20:  $r^{acceptant}$  trajectories for different ages

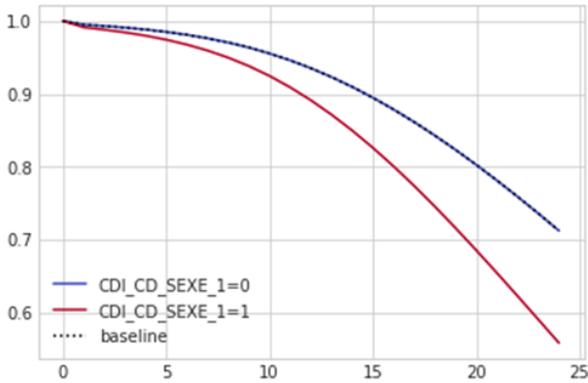


Figure 19:  $r^{acceptant}$  trajectories by gender

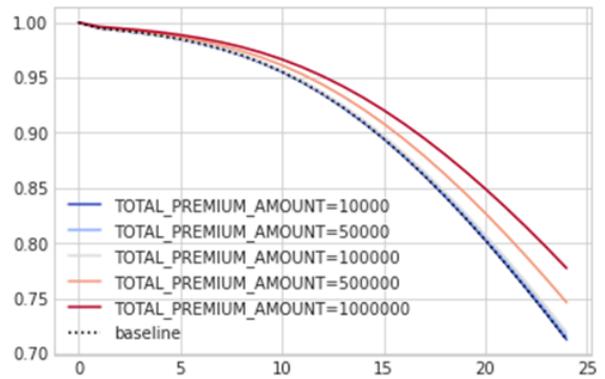


Figure 21:  $r^{acceptant}$  trajectories for different face amounts

**A.6.2. RSF**

We obtain better results than Cox in terms of concordance index at the cost of very high computation time for one training with one set of parameters - 5days without parallelisation, 4 hours with - compared to a few seconds for cox model.

Some of the results we obtain are displayed below:

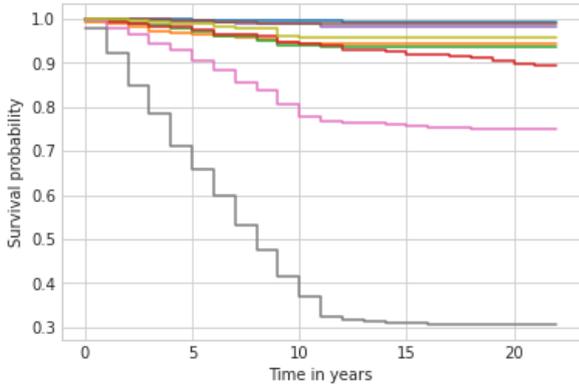


Figure 22: 10 policyholders' survival curve for  $r^{acceptant}$  with RSF

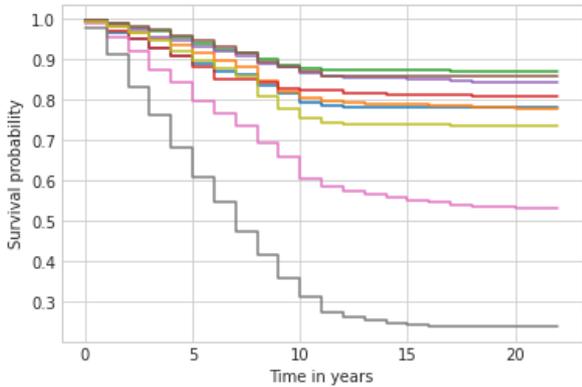


Figure 23: 10 policyholders' survival curve for  $r^{lasper}$  with RSF

Weight	Feature
0.3148 ± 0.0064	Age_souscription
0.0100 ± 0.0008	CDI_CD_SEXE_1
0.0091 ± 0.0014	PRODUIT_2
0.0077 ± 0.0006	TOTAL_PREMIUM_AMOUNT
0.0013 ± 0.0004	Nb_Contrats
0.0010 ± 0.0003	PRODUIT_3

Table 3: Covariates importance for  $r^{acceptant}$  with RSF

Weight	Feature
0.1838 ± 0.0045	Age_souscription
0.0415 ± 0.0018	TOTAL_PREMIUM_AMOUNT
0.0083 ± 0.0011	CDI_CD_SEXE_1
0.0026 ± 0.0013	PRODUIT_2
0.0022 ± 0.0006	PRODUIT_3
0.0020 ± 0.0006	Nb_Contrats

Table 4: Covariates importance for  $r^{lasper}$  with RSF

**A.6.3. XGSB**

We obtain better results than Cox and slightly better results than RSF in terms of concordance index at the cost of even higher computation time for one training with one set of parameters - 10h with great parallelisation - compared to a few seconds for Cox model.

Some of the results we obtain are displayed below:

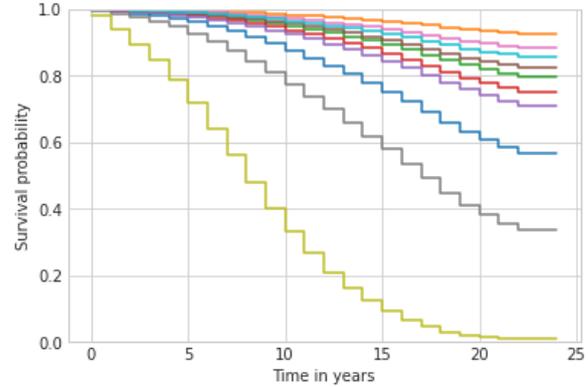


Figure 24: 10 policyholders' survival curve for  $r^{acceptant}$  with GBM

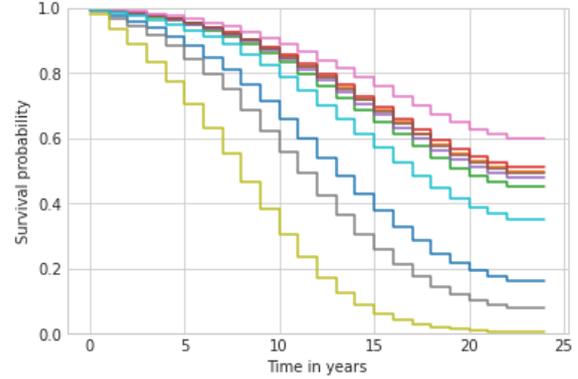


Figure 25: 10 policyholders' survival curve for  $r^{lasper}$  with GBM

Weight	Feature
0.3274 ± 0.0071	Age_souscription
0.0104 ± 0.0006	TOTAL_PREMIUM_AMOUNT
0.0100 ± 0.0008	CDI_CD_SEXE_1
0.0025 ± 0.0005	PRODUIT_2
0.0005 ± 0.0001	Nb_Contrats
0.0000 ± 0.0001	PRODUIT_3

Table 5: Covariates importance for  $r^{acceptant}$  with GBM

Weight	Feature
0.1872 ± 0.0039	Age_souscription
0.0438 ± 0.0020	TOTAL_PREMIUM_AMOUNT
0.0134 ± 0.0014	PRODUIT_2
0.0076 ± 0.0009	CDI_CD_SEXE_1
0.0051 ± 0.0006	PRODUIT_3
0.0011 ± 0.0004	Nb_Contrats

Table 6: Covariates importance for  $r^{lasper}$  with GBM

**A.6.4. Final survival model**

The final concordance index scores are displayed below:

	Concordance Index	
	$r^{lapses}$	$r^{acceptant}$
Cox model	69,5%	80,7%
RSF	71,6%	83,7%
GBSM	73,0%	84,1%

Table 7: Survival models comparison)

**A.7. Other LMS results**

LMS	p	$\delta$	$\gamma$	c	d	T
A-1	2,50%	0,04%	25%	10	1,50%	5
A-2	2,50%	0,04%	25%	10	1,50%	20
A-3	2,50%	0,04%	25%	100	1,50%	5
A-4	2,50%	0,04%	25%	100	1,50%	20
A-5	2,50%	0,04%	5%	10	1,50%	5
A-6	2,50%	0,04%	5%	10	1,50%	20
A-7	2,50%	0,04%	5%	100	1,50%	5
A-8	2,50%	0,04%	5%	100	1,50%	20
A-9	2,50%	0,10%	25%	10	1,50%	5
A-10	2,50%	0,10%	25%	10	1,50%	20
A-11	2,50%	0,10%	25%	100	1,50%	5
A-12	2,50%	0,10%	25%	100	1,50%	20
A-13	2,50%	0,10%	5%	10	1,50%	5
A-14	2,50%	0,10%	5%	10	1,50%	20
A-15	2,50%	0,10%	5%	100	1,50%	5
A-16	2,50%	0,10%	5%	100	1,50%	20
A-17	1,50%	0,04%	25%	10	1,50%	5
A-18	1,50%	0,04%	25%	10	1,50%	20
A-19	1,50%	0,04%	25%	100	1,50%	5
A-20	1,50%	0,04%	25%	100	1,50%	20
A-21	1,50%	0,04%	5%	10	1,50%	5
A-22	1,50%	0,04%	5%	10	1,50%	20
A-23	1,50%	0,04%	5%	100	1,50%	5
A-24	1,50%	0,04%	5%	100	1,50%	20
A-25	1,50%	0,10%	25%	10	1,50%	5
A-26	1,50%	0,10%	25%	10	1,50%	20
A-27	1,50%	0,10%	25%	100	1,50%	5
A-28	1,50%	0,10%	25%	100	1,50%	20
A-29	1,50%	0,10%	5%	10	1,50%	5
A-30	1,50%	0,10%	5%	10	1,50%	20
A-31	1,50%	0,10%	5%	100	1,50%	5
A-32	1,50%	0,10%	5%	100	1,50%	20
B-1	2,50%	0,08%	20%	10	1,50%	5
B-2	2,50%	0,08%	20%	10	1,50%	20
B-3	2,50%	0,08%	20%	100	1,50%	5
B-4	2,50%	0,08%	20%	100	1,50%	20
B-5	2,50%	0,08%	10%	10	1,50%	5
B-6	2,50%	0,08%	10%	10	1,50%	20
B-7	2,50%	0,08%	10%	100	1,50%	5
B-8	2,50%	0,08%	10%	100	1,50%	20
B-9	2,50%	0,20%	20%	10	1,50%	5
B-10	2,50%	0,20%	20%	10	1,50%	20

LMS	p	$\delta$	$\gamma$	c	d	T
B-11	2,50%	0,20%	20%	100	1,50%	5
B-12	2,50%	0,20%	20%	100	1,50%	20
B-13	2,50%	0,20%	10%	10	1,50%	5
B-14	2,50%	0,20%	10%	10	1,50%	20
B-15	2,50%	0,20%	10%	100	1,50%	5
B-16	2,50%	0,20%	10%	100	1,50%	20
B-17	1,50%	0,08%	20%	10	1,50%	5
B-18	1,50%	0,08%	20%	10	1,50%	20
B-19	1,50%	0,08%	20%	100	1,50%	5
B-20	1,50%	0,08%	20%	100	1,50%	20
B-21	1,50%	0,08%	10%	10	1,50%	5
B-22	1,50%	0,08%	10%	10	1,50%	20
B-23	1,50%	0,08%	10%	100	1,50%	5
B-24	1,50%	0,08%	10%	100	1,50%	20
B-25	1,50%	0,20%	20%	10	1,50%	5
B-26	1,50%	0,20%	20%	10	1,50%	20
B-27	1,50%	0,20%	20%	100	1,50%	5
B-28	1,50%	0,20%	20%	100	1,50%	20
B-29	1,50%	0,20%	10%	10	1,50%	5
B-30	1,50%	0,20%	10%	10	1,50%	20
B-31	1,50%	0,20%	10%	100	1,50%	5
B-32	1,50%	0,20%	10%	100	1,50%	20

Table 8: More LMS

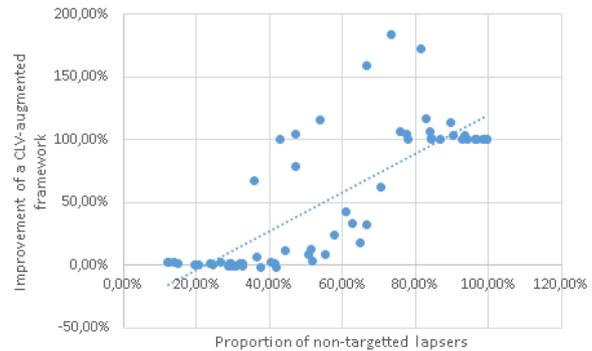


Figure 26: Correlation between the proportion of non-targetted lapsers and the improvement of a CLV-augmented LMS

CLV-based lapse management strategy

N°	time (s)	Model	% target diff	Accuracy		Retention gain		RG/target		Improvement
				$y_i$	$\tilde{y}_i$	$y_i$	$\tilde{y}_i$	$y_i$	$\tilde{y}_i$	
A-1	4949	CART	62,58%	92,3%	85,3%	114 661	219 655	4,48	38,20	91,57%
		RF		92,9%	85,4%	232 314	287 884	9,82	56,65	23,92%
		XGB		93,4%	85,8%	243 365	324 952	9,61	54,64	33,52%
A-2	6111	CART	26,66%	92,3%	89,8%	7 092 097	6 142 119	277,00	353,83	-13,39%
		RF		92,9%	90,2%	6 596 374	5 696 455	278,47	351,02	-13,64%
		XGB		93,4%	90,9%	7 308 721	7 432 688	288,92	404,84	1,70%
A-3	4603	CART	93,50%	92,3%	83,3%	- 2 187 622	- 8 224	- 85,52	- 31,09	99,62%
		RF		92,9%	83,4%	- 1 900 265	45 483	- 80,18	194,35	102,39%
		XGB		93,4%	83,5%	- 2 032 650	77 481	- 80,39	174,44	103,81%
A-4	5555	CART	55,37%	92,3%	86,5%	4 789 814	5 117 844	187,00	577,74	6,85%
		RF		92,9%	86,4%	4 463 796	4 255 175	188,47	566,05	-4,67%
		XGB		93,4%	86,8%	5 032 706	5 433 366	198,92	610,26	7,96%
A-5	4753	CART	86,72%	92,3%	83,6%	- 514 477	- 112 372	- 20,08	- 86,48	78,16%
		RF		92,9%	83,4%	- 323 544	- 3 937	- 13,65	- 28,28	98,78%
		XGB		93,4%	83,3%	- 383 004	-	- 15,14	-	100,00%
A-6	5803	CART	44,27%	92,3%	87,9%	335 810	517 224	13,17	39,91	54,02%
		RF		92,9%	87,9%	655 350	661 021	27,68	61,13	0,87%
		XGB		93,4%	88,6%	654 219	729 493	25,86	58,22	11,51%
A-7	4241	CART	99,09%	92,3%	83,3%	- 2 816 759	- 10 205	- 110,08	- 384,04	99,64%
		RF		92,9%	83,3%	- 2 456 122	1 013	- 103,65	66,30	100,04%
		XGB		93,4%	83,3%	- 2 659 020	243	- 105,14	15,92	100,01%
A-8	5164	CART	82,78%	92,3%	84,0%	- 1 966 473	- 46 323	- 76,83	- 22,31	97,64%
		RF		92,9%	84,0%	- 1 477 229	253 885	- 62,32	149,67	117,19%
		XGB		93,4%	84,1%	- 1 621 796	273 243	- 64,14	117,83	116,85%
A-9	4781	CART	77,60%	92,3%	83,7%	- 825 372	- 161 100	- 32,19	- 127,87	80,48%
		RF		92,9%	83,4%	- 384 736	8 596	- 16,22	32,12	102,23%
		XGB		93,4%	83,6%	- 498 263	22 337	- 19,70	35,47	104,48%
A-10	6075	CART	29,10%	92,3%	89,7%	4 614 513	4 483 831	180,36	266,33	-2,83%
		RF		92,9%	89,9%	4 973 929	4 328 724	210,01	280,90	-12,97%
		XGB		93,4%	90,7%	5 354 770	5 368 917	211,69	301,57	0,26%
A-11	4506	CART	96,56%	92,3%	83,2%	- 3 127 655	- 118 886	- 122,19	- 2 230,39	96,20%
		RF		92,9%	83,3%	- 2 517 315	1 340	- 106,22	87,71	100,05%
		XGB		93,4%	83,3%	- 2 774 278	736	- 109,70	52,00	100,03%
A-12	5534	CART	57,93%	92,3%	86,2%	2 312 231	3 310 314	90,36	412,71	43,17%
		RF		92,9%	86,1%	2 841 351	3 129 652	120,01	465,74	10,15%
		XGB		93,4%	86,6%	3 078 755	3 825 920	121,69	475,53	24,27%
A-13	4640	CART	92,91%	92,3%	83,3%	- 1 201 626	- 163 056	- 46,87	- 1 838,44	86,43%
		RF		92,9%	83,3%	- 717 620	- 5 339	- 30,28	- 354,24	99,26%
		XGB		93,4%	83,3%	- 875 378	508	- 34,60	16,26	100,06%
A-14	5739	CART	47,12%	92,3%	87,3%	- 1 476 651	- 831 019	- 57,49	- 77,99	43,72%
		RF		92,9%	86,0%	- 380 683	126 532	- 16,03	21,14	133,24%
		XGB		93,4%	85,5%	- 644 389	29 382	- 25,47	7,10	104,56%
A-15	4216	CART	99,61%	92,3%	83,3%	- 3 503 908	- 97 263	- 136,87	- 2 354,34	97,22%
		RF		92,9%	83,3%	- 2 850 198	-	- 120,28	-	100,00%
		XGB		93,4%	83,3%	- 3 151 393	-	- 124,60	-	100,00%
A-16	5096	CART	84,46%	92,3%	83,8%	- 3 778 933	- 734 773	- 147,49	- 418,58	80,56%
		RF		92,9%	83,5%	- 2 513 261	8 914	- 106,03	20,13	100,35%
		XGB		93,4%	83,6%	- 2 920 405	34 492	- 115,47	45,75	101,18%

CLV-based lapse management strategy

N°	time (s)	Model	% target diff	Accuracy		Retention gain		RG/target		Improvement
				$y_i$	$\tilde{y}_i$	$y_i$	$\tilde{y}_i$	$y_i$	$\tilde{y}_i$	
A-17	5390	CART	28,74%	92,3%	89,5%	5 100 456	4 899 479	199,11	279,88	-3,94%
		RF		92,9%	89,8%	4 635 482	4 226 648	195,69	276,06	-8,82%
		XGB		93,4%	90,2%	5 196 736	5 138 253	205,40	299,27	-1,13%
A-18	6452	CART	12,12%	92,3%	91,3%	52 090 240	47 706 070	2 034,15	2 170,64	-8,42%
		RF		92,9%	91,9%	46 171 160	42 049 900	1 949,05	2 082,36	-8,93%
		XGB		93,4%	92,5%	51 629 950	52 606 740	2 040,95	2 339,70	1,89%
A-19	4913	CART	64,89%	92,3%	85,2%	2 798 173	3 182 143	109,11	481,60	13,72%
		RF		92,9%	85,2%	2 502 903	2 743 070	105,69	554,76	9,60%
		XGB		93,4%	85,6%	2 920 720	3 438 303	115,40	576,64	17,72%
A-20	6160	CART	29,03%	92,3%	89,6%	49 787 960	45 366 730	1 944,15	2 616,32	-8,88%
		RF		92,9%	90,0%	44 038 580	39 947 830	1 859,05	2 547,89	-9,29%
		XGB		93,4%	90,6%	49 353 940	49 789 670	1 950,95	2 796,17	0,88%
A-21	5079	CART	51,69%	92,3%	86,8%	482 682	544 887	18,85	53,99	12,89%
		RF		92,9%	86,8%	557 090	554 195	23,52	65,17	-0,52%
		XGB		93,4%	87,1%	607 670	624 556	24,01	64,79	2,78%
A-22	6199	CART	23,94%	92,3%	90,2%	9 335 438	8 527 444	364,60	454,78	-8,66%
		RF		92,9%	90,6%	8 570 307	7 931 029	361,80	460,42	-7,46%
		XGB		93,4%	91,2%	9 518 466	9 581 934	376,27	501,56	0,67%
A-23	4601	CART	89,51%	92,3%	83,6%	- 1 819 600	135 305	- 71,15	121,80	107,44%
		RF		92,9%	83,5%	- 1 575 489	159 620	- 66,48	215,65	110,13%
		XGB		93,4%	83,7%	- 1 668 346	228 226	- 65,99	208,69	113,68%
A-24	5650	CART	50,83%	92,3%	87,0%	7 033 156	7 124 100	274,60	680,08	1,29%
		RF		92,9%	87,0%	6 437 729	6 364 477	271,80	711,89	-1,14%
		XGB		93,4%	87,4%	7 242 450	7 840 770	286,27	771,71	8,26%
A-25	5379	CART	30,97%	92,3%	89,2%	4 160 423	3 882 623	162,44	241,06	-6,68%
		RF		92,9%	89,5%	4 018 432	3 666 219	169,65	249,54	-8,76%
		XGB		93,4%	90,0%	4 455 108	4 410 629	176,09	267,87	-1,00%
A-26	6410	CART	12,52%	92,3%	91,3%	49 612 660	45 948 690	1 937,51	2 083,30	-7,39%
		RF		92,9%	91,9%	44 548 720	40 814 960	1 880,59	2 029,68	-8,38%
		XGB		93,4%	92,5%	49 676 000	50 549 740	1 963,72	2 260,20	1,76%
A-27	4887	CART	66,67%	92,3%	85,1%	1 858 140	2 575 538	72,44	442,86	38,61%
		RF		92,9%	85,0%	1 885 853	2 387 018	79,65	531,25	26,57%
		XGB		93,4%	85,4%	2 179 093	2 879 880	86,09	544,35	32,16%
A-28	6047	CART	29,42%	92,3%	89,4%	47 310 370	43 168 880	1 847,51	2 519,41	-8,75%
		RF		92,9%	89,9%	42 416 140	38 573 620	1 790,59	2 504,61	-9,06%
		XGB		93,4%	90,5%	47 399 990	47 812 830	1 873,72	2 721,63	0,87%
A-29	5070	CART	53,79%	92,3%	86,5%	- 204 467	- 5 098	- 7,95	- 1,66	97,51%
		RF		92,9%	86,1%	163 014	273 435	6,90	40,30	67,74%
		XGB		93,4%	86,8%	115 297	248 982	4,55	28,64	115,95%
A-30	6179	CART	24,36%	92,3%	90,3%	7 522 978	7 058 487	293,94	382,06	-6,17%
		RF		92,9%	90,6%	7 534 275	7 068 293	318,08	411,80	-6,18%
		XGB		93,4%	91,2%	8 219 857	8 265 167	324,94	442,88	0,55%
A-31	4627	CART	90,18%	92,3%	83,6%	- 2 506 749	- 139 983	- 97,95	- 121,44	94,42%
		RF		92,9%	83,5%	- 1 969 564	73 101	- 83,10	111,49	103,71%
		XGB		93,4%	83,6%	- 2 160 719	76 641	- 85,45	93,28	103,55%
A-32	5679	CART	51,25%	92,3%	86,8%	5 220 695	5 811 833	203,94	583,55	11,32%
		RF		92,9%	86,9%	5 401 696	5 269 505	228,08	605,69	-2,45%
		XGB		93,4%	87,4%	5 943 841	6 682 230	234,94	670,03	12,42%

CLV-based lapse management strategy

N°	time (s)	Model	% target diff	Accuracy		Retention gain		RG/target		Improvement
				$y_i$	$\tilde{y}_i$	$y_i$	$\tilde{y}_i$	$y_i$	$\tilde{y}_i$	
B-1	4778	CART	75,89%	92,3%	84,0%	- 627 165	- 148 913	- 24,46	- 65,19	76,26%
		RF		92,9%	83,7%	- 280 855	11 973	- 11,84	9,57	104,26%
		XGB		93,4%	84,1%	- 366 103	25 099	- 14,47	12,30	106,86%
B-2	6074	CART	29,70%	92,3%	89,7%	3 862 156	3 397 247	150,95	203,11	-12,04%
		RF		92,9%	89,9%	4 127 224	3 550 730	174,26	230,67	-13,97%
		XGB		93,4%	90,6%	4 451 686	4 408 819	175,99	250,17	-0,96%
B-3	4528	CART	96,60%	92,3%	83,2%	- 2 929 448	- 85 465	- 114,46	- 1 482,06	97,08%
		RF		92,9%	83,3%	- 2 413 433	3 724	- 101,84	- 108,33	100,15%
		XGB		93,4%	83,3%	- 2 642 119	9 092	- 104,47	93,79	100,34%
B-4	5476	CART	60,93%	92,3%	85,9%	1 559 874	2 471 262	60,95	329,63	58,43%
		RF		92,9%	85,8%	1 994 645	2 517 111	84,26	422,45	26,19%
		XGB		93,4%	86,3%	2 175 670	3 089 897	85,99	422,77	42,02%
B-5	4708	CART	84,33%	92,3%	83,4%	- 857 439	- 159 856	- 33,45	- 218,16	81,36%
		RF		92,9%	83,3%	- 484 459	40	- 20,44	7,23	100,01%
		XGB		93,4%	83,3%	- 596 203	897	- 23,57	46,96	100,15%
B-6	5906	CART	36,63%	92,3%	88,8%	705 721	922 490	27,69	60,21	30,72%
		RF		92,9%	88,9%	1 352 182	1 269 349	57,11	97,63	-6,13%
		XGB		93,4%	89,6%	1 342 882	1 428 722	53,09	96,76	6,39%
B-7	4400	CART	98,49%	92,3%	83,2%	- 3 159 722	- 39 633	- 123,45	- 1 230,61	98,75%
		RF		92,9%	83,3%	- 2 617 037	1 024	- 110,44	0,56	100,04%
		XGB		93,4%	83,3%	- 2 872 219	295	- 113,57	19,31	100,01%
B-8	5278	CART	73,18%	92,3%	84,6%	- 1 596 562	169 852	- 62,31	41,78	110,64%
		RF		92,9%	84,6%	- 780 396	637 625	- 32,89	194,52	181,71%
		XGB		93,4%	85,0%	- 933 133	780 845	- 36,91	188,79	183,68%
B-9	4601	CART	94,12%	92,3%	83,3%	- 2 380 789	- 113 444	- 92,86	- 840,25	95,24%
		RF		92,9%	83,3%	- 1 403 468	317	- 59,21	7,96	100,02%
		XGB		93,4%	83,3%	- 1 724 731	3 980	- 68,17	149,44	100,23%
B-10	5947	CART	35,98%	92,3%	89,0%	- 760 449	429 196	- 29,35	29,80	156,44%
		RF		92,9%	88,5%	1 175 540	1 354 131	49,71	118,11	15,19%
		XGB		93,4%	89,8%	871 455	1 456 080	34,48	96,25	67,09%
B-11	4229	CART	99,16%	92,3%	83,3%	- 4 683 072	- 48 985	- 182,86	- 1 186,22	98,95%
		RF		92,9%	83,3%	- 3 536 046	-	- 149,21	-	100,00%
		XGB		93,4%	83,3%	- 4 000 747	-	- 158,17	-	100,00%
B-12	5391	CART	66,76%	92,3%	85,0%	- 3 062 732	- 388 289	- 119,35	- 80,44	87,32%
		RF		92,9%	84,7%	- 957 039	710 688	- 40,29	220,55	174,26%
		XGB		93,4%	85,3%	- 1 404 561	834 198	- 55,52	163,88	159,39%
B-13	4493	CART	96,30%	92,3%	83,3%	- 2 358 179	- 159 922	- 91,98	- 2 793,13	93,22%
		RF		92,9%	83,3%	- 1 384 098	-	- 58,40	-	100,00%
		XGB		93,4%	83,3%	- 1 705 577	-	- 67,42	-	100,00%
B-14	5851	CART	42,98%	92,3%	87,8%	- 3 251 762	- 1 761 821	- 126,63	- 143,20	45,82%
		RF		92,9%	86,4%	- 1 013 089	79 273	- 42,69	11,90	107,82%
		XGB		93,4%	83,3%	- 1 582 006	4 396	- 62,52	287,68	100,28%
B-15	4040	CART	99,67%	92,3%	83,3%	- 4 660 462	- 38 969	- 181,98	- 2 075,03	99,16%
		RF		92,9%	83,3%	- 3 516 676	-	- 148,40	-	100,00%
		XGB		93,4%	83,3%	- 3 981 592	161	- 157,42	10,53	100,00%
B-16	5182	CART	77,97%	92,3%	84,2%	- 5 554 044	- 1 491 522	- 216,63	- 549,23	73,15%
		RF		92,9%	83,6%	- 3 145 668	52 475	- 132,69	84,54	101,67%
		XGB		93,4%	83,3%	- 3 858 022	-	- 152,52	-	100,00%

CLV-based lapse management strategy

N°	time (s)	Model	% target diff	Accuracy		Retention gain		RG/target		Improvement
				$y_i$	$\tilde{y}_i$	$y_i$	$\tilde{y}_i$	$y_i$	$\tilde{y}_i$	
B-17	5324	CART	32,66%	92,3%	88,9%	3 361 471	3 037 200	131,25	191,31	-9,65%
		RF		92,9%	89,3%	3 241 680	2 911 023	136,86	204,43	-10,20%
		XGB		93,4%	89,6%	3 596 593	3 546 671	142,15	222,04	-1,39%
B-18	6411	CART	13,83%	92,3%	91,1%	39 860 670	37 695 680	1 556,66	1 778,71	-5,43%
		RF		92,9%	91,7%	35 787 050	32 345 100	1 510,72	1 654,32	-9,62%
		XGB		93,4%	92,0%	39 908 670	40 886 810	1 577,61	1 848,71	2,45%
B-19	4853	CART	70,34%	92,3%	84,7%	1 059 189	1 813 631	41,25	392,14	71,23%
		RF		92,9%	84,8%	1 109 101	1 808 616	46,86	474,33	63,07%
		XGB		93,4%	85,0%	1 320 578	2 141 271	52,15	482,34	62,15%
B-20	5973	CART	31,76%	92,3%	89,2%	37 558 390	34 068 550	1 466,66	2 125,97	-9,29%
		RF		92,9%	89,4%	33 654 470	30 032 580	1 420,72	2 072,47	-10,76%
		XGB		93,4%	90,1%	37 632 650	38 008 480	1 487,61	2 277,17	1,00%
B-21	5228	CART	41,79%	92,3%	87,7%	1 136 879	1 179 837	44,40	92,50	3,78%
		RF		92,9%	88,1%	1 276 808	1 188 256	53,91	104,81	-6,94%
		XGB		93,4%	88,7%	1 385 145	1 356 864	54,74	104,76	-2,04%
B-22	6296	CART	19,52%	92,3%	90,7%	18 704 980	17 177 190	730,55	852,81	-8,17%
		RF		92,9%	91,1%	17 182 100	15 732 340	725,34	859,29	-8,44%
		XGB		93,4%	91,5%	19 071 370	19 050 020	753,90	939,00	-0,11%
B-23	4746	CART	81,36%	92,3%	84,1%	- 1 165 404	458 223	- 45,60	172,83	139,32%
		RF		92,9%	84,0%	- 855 770	525 335	- 36,09	288,55	161,39%
		XGB		93,4%	84,1%	- 890 871	645 445	- 35,26	310,86	172,45%
B-24	5845	CART	40,47%	92,3%	88,2%	16 402 700	15 013 310	640,55	1 093,43	-8,47%
		RF		92,9%	88,4%	15 049 520	13 423 040	635,34	1 122,81	-10,81%
		XGB		93,4%	88,9%	16 795 360	17 144 260	663,90	1 247,50	2,08%
B-25	5274	CART	37,42%	92,3%	88,6%	1 607 847	1 839 864	62,84	126,33	14,43%
		RF		92,9%	88,7%	2 119 067	1 923 982	89,49	152,71	-9,21%
		XGB		93,4%	89,2%	2 237 965	2 194 469	88,45	155,54	-1,94%
B-26	6425	CART	14,83%	92,3%	91,1%	35 238 060	32 690 970	1 376,37	1 558,26	-7,23%
		RF		92,9%	91,6%	32 835 370	29 986 540	1 386,17	1 543,12	-8,68%
		XGB		93,4%	92,0%	36 328 440	36 803 630	1 436,10	1 688,53	1,31%
B-27	4811	CART	73,92%	92,3%	84,3%	- 694 436	751 404	- 27,16	226,99	208,20%
		RF		92,9%	84,4%	- 13 512	1 018 369	- 0,51	356,48	7636,98%
		XGB		93,4%	84,7%	- 38 050	1 253 252	- 1,55	345,94	3393,68%
B-28	5995	CART	32,61%	92,3%	89,1%	32 935 780	29 342 930	1 286,37	1 847,71	-10,91%
		RF		92,9%	89,4%	30 702 790	27 725 620	1 296,17	1 933,38	-9,70%
		XGB		93,4%	90,0%	34 052 420	34 390 060	1 346,10	2 094,90	0,99%
B-29	5143	CART	47,03%	92,3%	87,3%	- 363 861	55 985	- 14,12	3,38	115,39%
		RF		92,9%	87,4%	377 170	488 284	15,95	49,62	29,46%
		XGB		93,4%	88,0%	275 772	491 567	10,89	44,89	78,25%
B-30	6243	CART	20,47%	92,3%	90,7%	14 747 500	13 838 380	576,23	690,22	-6,16%
		RF		92,9%	91,1%	14 816 830	13 378 460	625,54	743,34	-9,71%
		XGB		93,4%	91,5%	16 146 490	16 169 440	638,30	814,80	0,14%
B-31	4730	CART	83,83%	92,3%	83,7%	- 2 666 144	- 487 716	- 104,12	- 267,75	81,71%
		RF		92,9%	83,7%	- 1 755 409	139 545	- 74,05	102,66	107,95%
		XGB		93,4%	83,7%	- 2 000 244	134 199	- 79,11	130,13	106,71%
B-32	5865	CART	41,41%	92,3%	88,0%	12 445 210	11 693 070	486,23	884,49	-6,04%
		RF		92,9%	88,3%	12 684 250	11 381 260	535,54	971,28	-10,27%
		XGB		93,4%	88,8%	13 870 470	14 101 470	548,30	1 048,38	1,67%

## References

- Akaike, H., 1973. Information theory and an extension of the maximum likelihood principle ». dans Second International Symposium on Information Theory , 267–281.
- Amit, Y., Geman, D., 1997. Shape quantization and recognition with randomized trees. *Neural Computation* 9, 1545–1588. doi:10.1162/neco.1997.9.7.1545.
- Ascarza, E., Neslin, S.A., Netzer, O., Anderson, Z., Fader, P.S., Gupta, S., Hardie, B., Lemmens, A., Libai, B., Neal, D.T., Provost, F., Schrift, R., 2018. In pursuit of enhanced customer retention management: Review, key issues, and future directions. Special Issue on 2016 Choice Symposium. *Customer Needs and Solutions* 5.
- Azzone, M., Barucci, E., Giuffrè Moncayo, G., Marazzina, D., 2022. A machine learning model for lapse prediction in life insurance contracts. *Expert Systems with Applications* 191, 116261. URL: <https://www.sciencedirect.com/science/article/pii/S0957417421015700>, doi:<https://doi.org/10.1016/j.eswa.2021.116261>.
- Bou-Hamad, I., Laroque, D., Ben-Ameur, H., 2011. A review of survival trees. *Statistics Surveys* 5, 44 – 71. URL: <https://doi.org/10.1214/09-SS047>, doi:10.1214/09-SS047.
- Breiman, L., 1996. Bagging predictors. *Mach. Learn.* 24, 123–140. URL: <https://doi.org/10.1023/A:1018054314350>, doi:10.1023/A:1018054314350.
- Breiman, L., 2001. Random forests. *Machine Learning* 45, 5–32. URL: <http://dx.doi.org/10.1023/A%3A1010933404324>, doi:10.1023/A:1010933404324.
- Breiman, L., Friedman, J., Stone, C., Olshen, R., 1984. *Classification and Regression Trees*. Taylor & Francis. URL: <https://books.google.fr/books?id=JwQx-W0mSyQc>.
- Buchardt, K., 2014. Dependent interest and transition rates in life insurance. *Insurance: Mathematics and Economics* 55. doi:10.1016/j.insmatheco.2014.01.004.
- Buchardt, K., Møller, T., Schmidt, K.B., 2015. Cash flows and policyholder behaviour in the semi-markov life insurance setup. *Scandinavian Actuarial Journal* 2015, 660–688. URL: <https://doi.org/10.1080/03461238.2013.879919>, doi:10.1080/03461238.2013.879919, arXiv:<https://doi.org/10.1080/03461238.2013.879919>.
- Chen, T., Guestrin, C., 2016. XGBoost: A scalable tree boosting system, in: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM, New York, NY, USA, pp. 785–794. URL: <http://doi.acm.org/10.1145/2939672.2939785>, doi:10.1145/2939672.2939785.
- Cox, Lin, . Annuity Lapse Modeling: Tobit or not Tobit ? *Society of Actuaries*.
- Cox, D., . Regression models and life-tables. *Journal of the Royal Statistical Society, Series B (Methodological)* 34, 187–220. URL: <http://www.jstor.org/stable/2985181>. jSTOR.
- Dar, A., Dodds, C., 1989. Interest rates, the emergency fund hypothesis and saving through endowment policies: Some empirical evidence for the u.k. *Journal of Risk and Insurance* 56, 415.
- Davidson-Pilon, C., 2019. lifelines: survival analysis in python. *Journal of Open Source Software* 4, 1317. URL: <https://doi.org/10.21105/joss.01317>, doi:10.21105/joss.01317.
- Donkers, B., Verhoef, P., Jong, M., 2007. Modeling clv: A test of competing models in the insurance industry. *Quantitative Marketing and Economics (QME)* 5, 163–190. URL: <https://EconPapers.repec.org/RePEc:kap:qmktec:v:5:y:2007:i:2:p:163-190>.
- Eling, M., Kiesenbauer, D., 2014. What policy features determine life insurance lapse? an analysis of the german market. *The Journal of Risk and Insurance* 81, 241–269. URL: <http://www.jstor.org/stable/24546804>.
- Eling, M., Kochanski, M., 2013. Research on lapse in life insurance: what has been done and what needs to be done? *Journal of Risk Finance* 14, 392–413. URL: <https://EconPapers.repec.org/RePEc:eme:jrfpps:v:14:y:2013:i:4:p:392-413>.
- Freund, Y., Schapire, R.E., 1996. Experiments with a new boosting algorithm, in: *International Conference on Machine Learning*, pp. 148–156. URL: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.51.6252>.
- Gatzert, N., Schmeiser, H., 2008. Assessing the risk potential of premium payment options in participating life insurance contracts. *The Journal of Risk and Insurance* 75, 691–712. URL: <http://www.jstor.org/stable/25145301>.
- Gemmo, I., Götz, M., 2016. Life insurance and demographic change: An empirical analysis of surrender decisions based on panel data. *ICIR Working Paper Series 24/16*. Goethe University Frankfurt, International Center for Insurance Regulation (ICIR). URL: <https://ideas.repec.org/p/zbw/icirwp/2416.html>.
- Grinsztajn, L., Oyallon, E., Varoquaux, G., 2022. Why do tree-based models still outperform deep learning on tabular data? URL: <https://arxiv.org/abs/2207.08815>, doi:10.48550/ARXIV.2207.08815.
- Gupta, S., 2009. Customer-based valuation. *Journal of Interactive Marketing* 23, 169–178. URL: <https://www.sciencedirect.com/science/article/pii/S109499680900036X>, doi:<https://doi.org/10.1016/j.intmar.2009.02.006>. anniversary Issue.
- Gupta, S., Hanssens, D., Hardie, B., Kahn, W., Kumar, V., Lin, N., Ravishanker, N., Sriram, S., 2006. Modeling customer lifetime value. *Journal of Service Research - J SERV RES* 9, 139–155. doi:10.1177/1094670506293810.
- Gupta, S., Lehmann, D.R., 2006. Customer lifetime value and firm valuation. *Journal of Relationship Marketing* 5, 87–110. URL: [https://doi.org/10.1300/J366v05n02\\_06](https://doi.org/10.1300/J366v05n02_06), doi:10.1300/J366v05n02\_06, arXiv:[https://doi.org/10.1300/J366v05n02\\_06](https://doi.org/10.1300/J366v05n02_06).
- Hanley, J., McNeil, B., 1982. The meaning and use of the area under a receiver operating characteristic (roc) curve. *Radiology* Apr;143(1):29–36. doi:10.1148/radiology.143.1.7063747.. PMID: 7063747.
- Harrell, F.E., Califf, R.M., Pryor, D.B., Lee, K.L., Rosati, R.A., 1982. Evaluating the yield of medical tests. *Jama* 247, 2543–2546.
- Hu, S., O’Hagan, A., Sweeney, J., Ghahramani, M., 2021. A spatial machine learning model for analysing customers’ lapse behaviour in life insurance. *Annals of Actuarial Science* 15, 367–393. doi:10.1017/S1748499520000329.
- Hwang, Y., Chan, L.F.S., Tsai, C.J., 2022. On voluntary terminations of life insurance: Differentiating surrender propensity from lapse propensity across product types. *North American Actuarial Journal* 26, 252–282. URL: <https://doi.org/10.1080/10920277.2021.1973507>, doi:10.1080/10920277.2021.1973507, arXiv:<https://doi.org/10.1080/10920277.2021.1973507>.
- Ishwaran, H., Gerds, T., Kogalur, U., Moore, R., Gange, S., Lau, B., . Random survival forests for competing risks. *Biostatistics* Oct;15(4):757–73. doi:10.1093/biostatistics/kxu010.. epub 2014 Apr 11. PMID: 24728979 ; PMID: PMC4173102.
- Ishwaran, H., Kogalur, U., 2007. Random survival forests for r. *R News* 7, 25–31. URL: <https://CRAN.R-project.org/doc/Rnews/>.
- Ishwaran, H., Kogalur, U.B., Blackstone, E.H., Lauer, M.S., 2008. Random survival forests. *The Annals of Applied Statistics* 2, 841 – 860. URL: <https://doi.org/10.1214/08-AOAS169>, doi:10.1214/08-AOAS169.
- Kagaoka, Y., 2005. Modeling insurance surrenders by the negative binomial model. *JAFEE International Conference* .
- Kiesenbauer, D., 2012. Main determinants of lapse in the german life insurance industry. *North American Actuarial Journal* 16, 52–73. URL: <https://doi.org/10.1080/10920277.2012.10590632>, doi:10.1080/10920277.2012.10590632, arXiv:<https://doi.org/10.1080/10920277.2012.10590632>.
- Kim, C., 2005. Modeling surrender and lapse rates with economic variables. *North American Actuarial Journal* 9, 56–70. URL: <https://doi.org/10.1080/10920277.2005.10596225>, doi:10.1080/10920277.2005.10596225, arXiv:<https://doi.org/10.1080/10920277.2005.10596225>.
- KPMG, 2020. First impressions: Ifrs 17 insurance contracts (2020 edition). URL: <https://assets.kpmg/content/dam/kpmg/ie/pdf/2020/09/ie-ifrs-17-first-impressions.pdf>.
- Kuo, W., Tsai, C., Chen, W.K., 2003. An empirical study on the lapse rate: The cointegration approach. *The Journal of Risk and Insurance* 70, 489–508. URL: <http://www.jstor.org/stable/3519905>.
- LeBlanc, M., Crowley, J., 1993. Survival trees by goodness of split. *Journal of the American Statistical Association* 88, 457. doi:10.2307/2290325.

- Lemmens, A., Gupta, S., 2020. Managing Churn to Maximize Profits. *Marketing Science* 39, 956–973. URL: <https://ideas.repec.org/a/inm/ormksc/v39y2020i5p956-973.html>, doi:10.1287/mksc.2020.1229.
- Loisel, S., Piette, P., Tsai, C.H.J., 2021. Applying economic measures to lapse risk management with machine learning approaches. *ASTIN Bulletin* 51, 839–871. doi:10.1017/asb.2021.10.
- Mantel, N., 1966. Evaluation of survival data and two new rank order statistics arising in its consideration. *Cancer Chemotherapy Reports. Part 1* 50, 163–170.
- Milhaud, X., Dutang, C., 2018. Lapse tables for lapse risk management in insurance: a competing risk approach. *European Actuarial Journal* 8, 97–126. URL: <https://hal.archives-ouvertes.fr/hal-01985256>, doi:10.1007/s13385-018-0165-7.
- Milhaud, X., Loisel, S., Maume-Deschamps, V., 2011. Surrender triggers in Life Insurance: what main features affect the surrender behavior in a classical economic context ? *Bulletin Français d'Actuariat* 11, 5–48. URL: <https://hal.archives-ouvertes.fr/hal-01985261>.
- Nolte, S., Schneider, J.C., 2017. Don't lapse into temptation: a behavioral explanation for policy surrender. *Journal of Banking & Finance* 79, 12–27. URL: <https://EconPapers.repec.org/RePEc:eee:jbfina:v:79:y:2017:i:c:p:12-27>.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., et al., 2011. Scikit-learn: Machine learning in python. *Journal of machine learning research* 12, 2825–2830.
- Pölsterl, S., 2020. scikit-survival: A library for time-to-event analysis built on top of scikit-learn. *Journal of Machine Learning Research* 21, 1–6. URL: <http://jmlr.org/papers/v21/20-729.html>.
- Pouffinas, T., Michaelide, G., 2018. Determinants of life insurance policy surrenders. *Modern Economy* 09, 1400–1422. doi:10.4236/me.2018.98089.
- Putter, H., Schumacher, M., Houwelingen, H., . On the relation between the cause-specific hazard and the subdistribution rate for competing risks data: The fine-gray model revisited. *Biom J*.
- Renshaw, A.E., Haberman, S., 1986. Statistical analysis of life assurance lapses. *Journal of the Institute of Actuaries* 113, 459–497.
- Routh, P., Roy, A., Meyer, J., 2021. Estimating customer churn under competing risks. *Journal of the Operational Research Society* 72, 1138–1155. URL: <https://doi.org/10.1080/01605682.2020.1776166>, doi:10.1080/01605682.2020.1776166, arXiv:<https://doi.org/10.1080/01605682.2020.1776166>.
- Russell, D.T., Fier, S.G., Carson, J.M., Dumm, R.E., 2013. An empirical analysis of life insurance policy surrender activity. *Journal of Insurance Issues* 36, 35–57. URL: <http://www.jstor.org/stable/41946336>.
- Shamsuddin, S., Noriszura, I., Roslan, N., 2022. What we know about research on life insurance lapse: A bibliometric analysis. *Risks* 10, 97. doi:10.3390/risks10050097.
- Sirak, A.S., 2015. Income and unemployment effects on life insurance lapse. Retrieved September 18, 2020.
- Vasudev, M., Bajaj, R., Escolano, A.A., . On the Drivers of Lapse Rates in Life Insurance. Sarjana thesis,. University of Barcelona. Barcelona, Spain.
- von Mutius, B., Huchzermeier, A., 2021. Customized targeting strategies for category coupons to maximize clv and minimize cost. *Journal of Retailing* 97, 764–779. URL: <https://www.sciencedirect.com/science/article/pii/S002243592100004X>, doi:<https://doi.org/10.1016/j.jretai.2021.01.004>. sl: Metrics and Analytics.
- Yu, L., Cheng, J., Lin, T., 2019. Life insurance lapse behaviour: evidence from China. *The Geneva Papers on Risk and Insurance - Issues and Practice* 44, 653–678. URL: [https://ideas.repec.org/a/pal/gpprii/v44y2019i4d10.1057\\_s41288-018-0104-5.html](https://ideas.repec.org/a/pal/gpprii/v44y2019i4d10.1057_s41288-018-0104-5.html), doi:10.1057/s41288-018-0104-5.
- Ćurak, M., Podrug, D., Poposki, K., 2015. Policyholder and insurance policy features as determinants of life insurance lapse - evidence from croatia. *Economics and Business Review* 1 (15), 58–77. doi:10.18559/ebr.2015.3.5.