Next Generation Underwriting

TING DAI, FSA, CERA

Head of APAC Data Analytics, SCOR Global Life

28 September 2020

SOCIETY OF ACTUARIES®

SOA Antitrust Compliance Guidelines

Active participation in the Society of Actuaries is an important aspect of membership. While the positive contributions of professional societies and associations are well-recognized and encouraged, association activities are vulnerable to close antitrust scrutiny. By their very nature, associations bring together industry competitors and other market participants.

The United States antitrust laws aim to protect consumers by preserving the free economy and prohibiting anti-competitive business practices; they promote competition. There are both state and federal antitrust laws, although state antitrust laws closely follow federal law. The Sherman Act, is the primary U.S. antitrust law pertaining to association activities. The Sherman Act prohibits every contract, combination or conspiracy that places an unreasonable restraint on trade. There are, however, some activities that are illegal under all circumstances, such as price fixing, market allocation and collusive bidding.

There is no safe harbor under the antitrust law for professional association activities. Therefore, association meeting participants should refrain from discussing any activity that could potentially be construed as having an anti-competitive effect. Discussions relating to product or service pricing, market allocations, membership restrictions, product standardization or other conditions on trade could arguably be perceived as a restraint on trade and may expose the SOA and its members to antitrust enforcement procedures.

While participating in all SOA in person meetings, webinars, teleconferences or side discussions, you should avoid discussing competitively sensitive information with competitors and follow these guidelines:

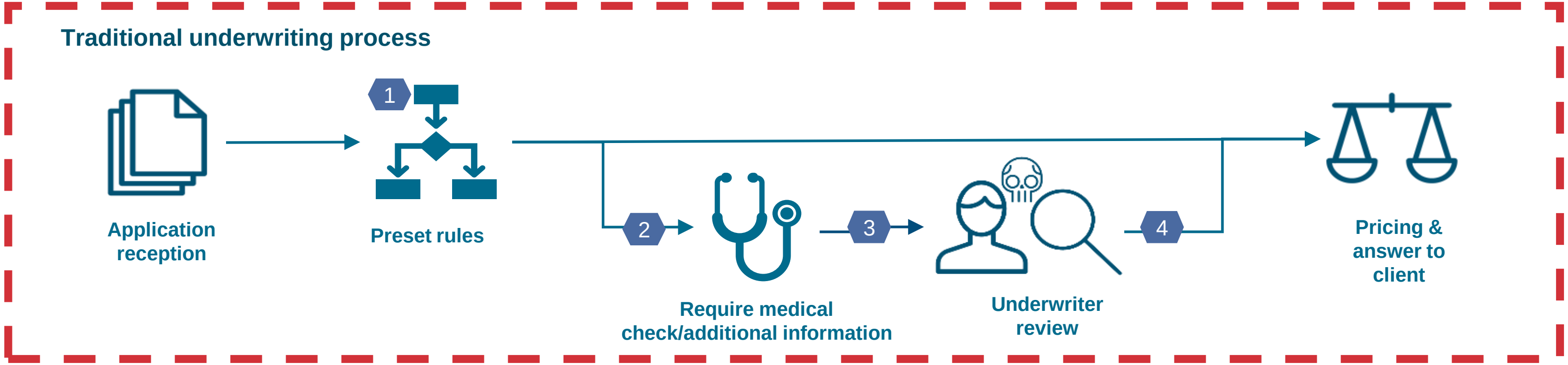
- -Do not discuss prices for services or products or anything else that might affect prices
- -Do not discuss what you or other entities plan to do in a particular geographic or product markets or with particular customers.
- -Do not speak on behalf of the SOA or any of its committees unless specifically authorized to do so.
- -Do leave a meeting where any anticompetitive pricing or market allocation discussion occurs.
- -Do alert SOA staff and/or legal counsel to any concerning discussions
- -Do consult with legal counsel before raising any matter or making a statement that may involve competitively sensitive information.

Adherence to these guidelines involves not only avoidance of antitrust violations, but avoidance of behavior which might be so construed. These guidelines only provide an overview of prohibited activities. SOA legal counsel reviews meeting agenda and materials as deemed appropriate and any discussion that departs from the formal agenda should be scrutinized carefully. Antitrust compliance is everyone’s responsibility; however, please seek legal counsel if you have any questions or concerns.

Presentation Disclaimer

Presentations are intended for educational purposes only and do not replace independent professional judgment. Statements of fact and opinions expressed are those of the participants individually and, unless expressly stated to the contrary, are not the opinion or position of the Society of Actuaries, its cosponsors or its committees. The Society of Actuaries does not endorse or approve, and assumes no responsibility for, the content, accuracy or completeness of the information presented. Attendees should note that the sessions are audio-recorded and may be published in various media, including print, audio and video formats without further notice.

Some pain points of traditional underwriting process

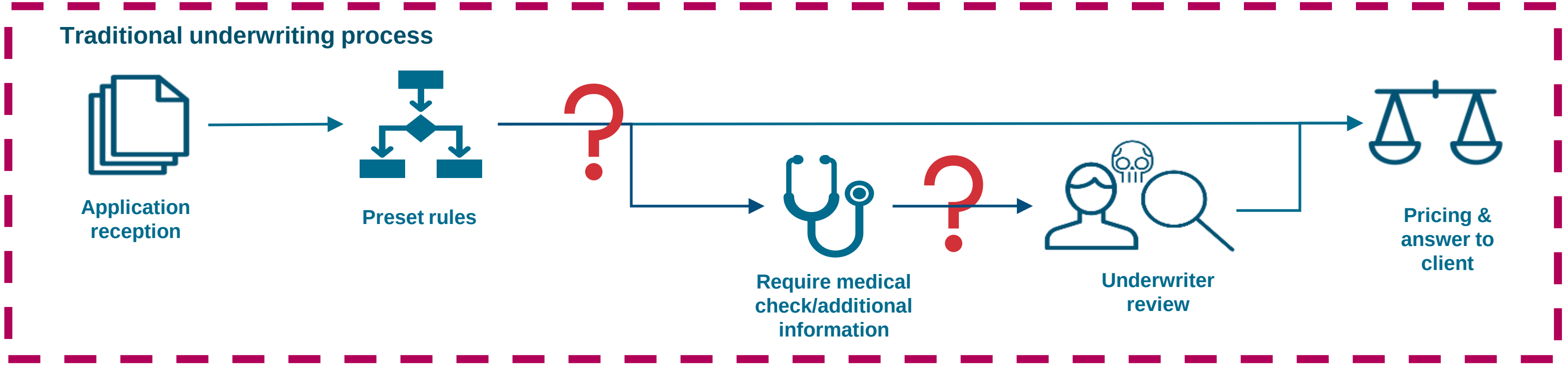


- Some pain points of the traditional underwriting process

1. The preset underwriting table is too rigid to reflect true risks, increase overall costs for insurers
2. Insurers ask clients to submit too many documents, increase clients unsatisfaction especially when clients can not benefit from providing documents
3. Short timeline, heavy workload for underwriters, no prioritization
4. Linking pricing and underwriting is not straightforward

USE CASE 1: Predictive Underwriting for Medex

Define the problem

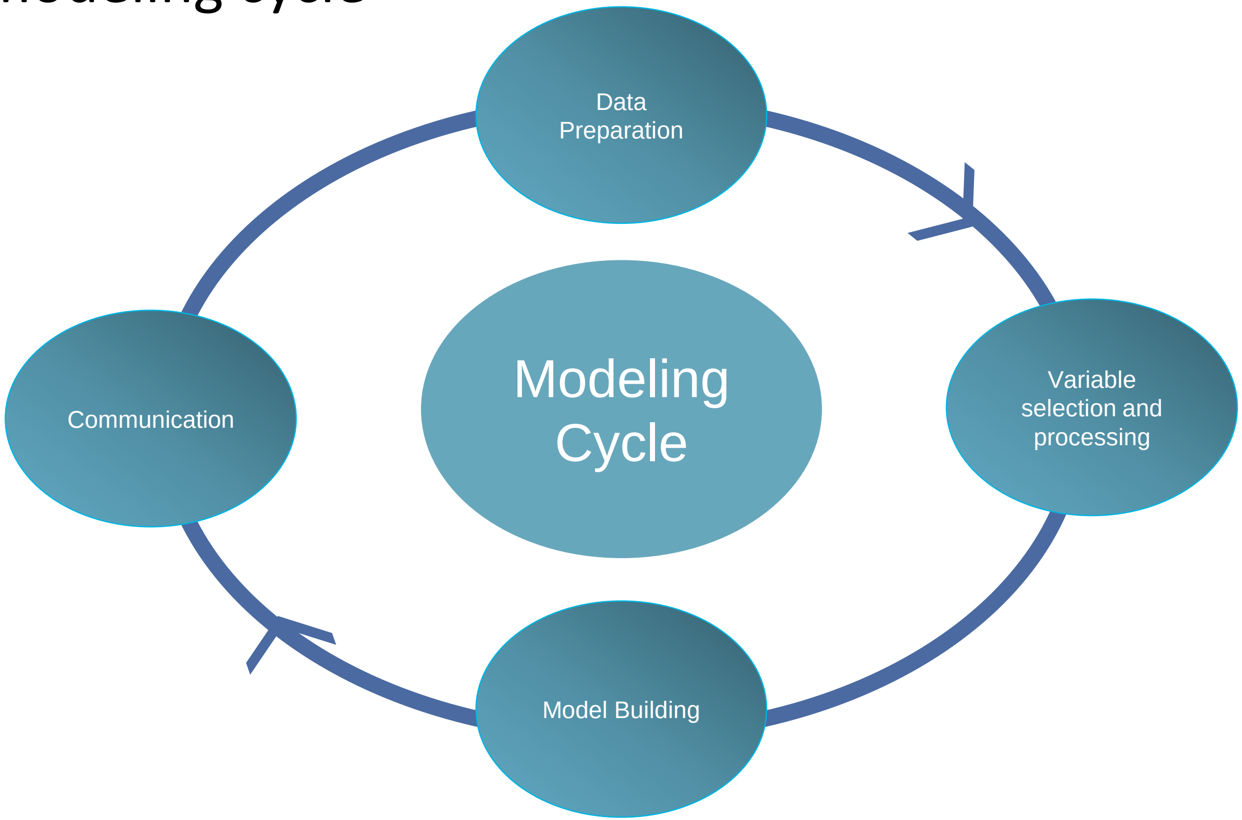


- The sales of medical insurance is booming in Asia. Our client is a major player in medical insurance.
- According to their underwriting table, all people above X years old need to go through medical underwriting. The cost is huge, and the sales process is not smooth
- For those below X years old, the underwriters will randomly choose some clients and ask them to go through medical underwriting. Due to resources limit, the selection should become more targeted.

Potential data resource

- In big data era, tons of data were created everyday
  - In 2015, no less than 2.5 quintillions of data created daily
  - Not just greater volume of data but also being recorded in new ways (search engines, social media, mobile devices)
  - But many insurance companies think that they do not have enough data comparing with tech firm. Is this the truth?
- The insurance has the most relevant data to insurance industry.
- For the underwriting models, more than 200 variables are extracted
  - Current underwriting decision
  - Claim experience
  - Policy application information
  - Past insurance purchase history
  - Past insurance claim history
  - Agent information
  - Other derived data

Typical modeling cycle

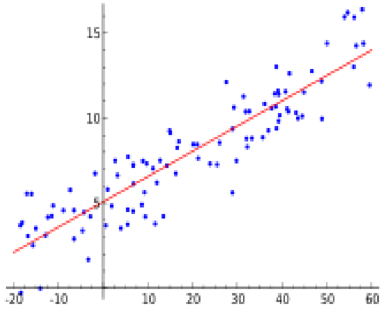


What algorithms to be used

GLM: Generalized Linear Model

Principle of the method

Model initially associated with statistics. In machine learning, few assumptions are made about the distribution of data. Only the parameterized model is then retained: $\mathrm{E}[Y|X] = g(\beta^t X)$



Advantages / Disadvantages

+

- Easy to interpret thanks to the parameterization

−

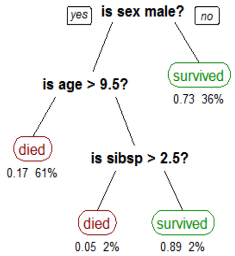
- Does not take into account the effects of non-linearity

CART: Classification and Regression Tree

Principle of the method

The algorithm allows to define a decision tree. At each node a variable is selected (according to a criterion to be defined by the user - variance, Gini) in order to determine two classes as homogeneous as possible

Example of a tree to describe the survivors of the Titanic



Advantages / Disadvantages

+

- Simple interpretation thanks to easy reading of the results
- Can handle large cases

−

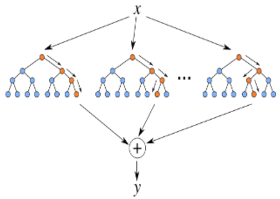
- Not very robust (requires boosting but loss of interpretation in Random Forest)

Ensemble: Bagging or Boosting

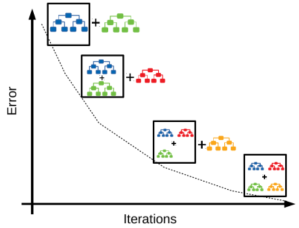
Principle of the method

Bagging: For a number of times N defined by the user, the database is resampled (discounted print run). On each new base, a CART is calibrated. The average prediction of N models provides the final prediction.

Boosting: Iterative method based on the weak learner notion. The idea is to sum different trees in order to decrease the learning error.



Bagging



Boosting

Advantages / Disadvantages

+

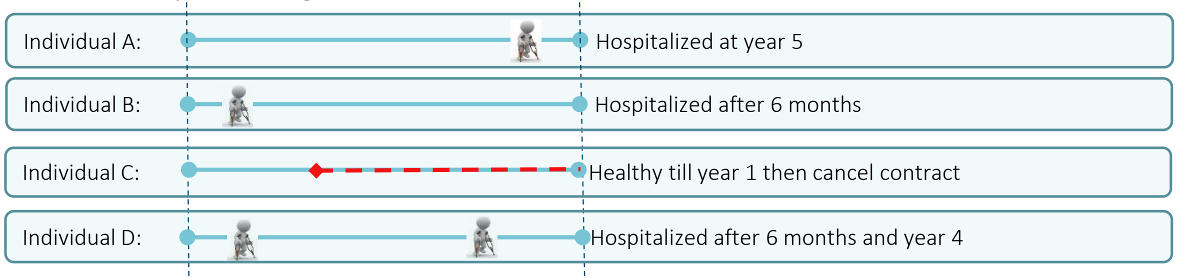
- Robust design
- Bagging: Easily parallelizable
- Boosting: Good generalization capacity and predictive power on tabular data

−

- Black Box, need other tool to correctly control the model
- Loss of interpretation

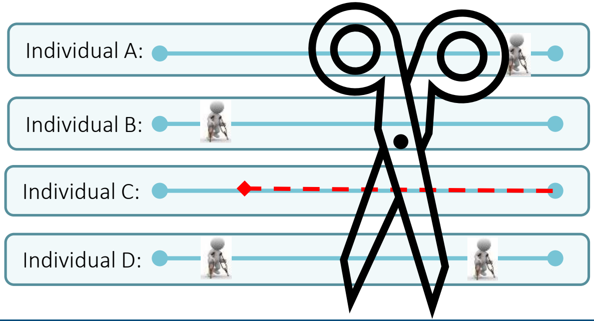
How time should be considered: Data scientist vs Actuaries

- Event to predict: whether an individual will be hospitalized
  - Is this definition of problem good?



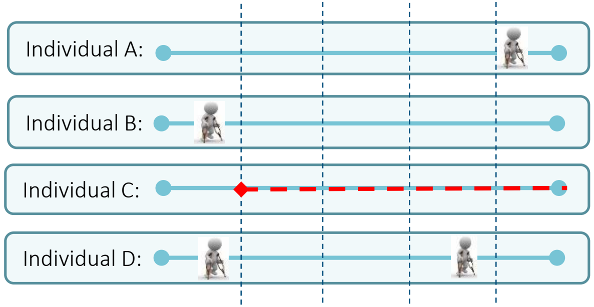
A typical data scientists' approach

Transform to classic classification problem: whether an individual will be hospitalized within 2 years



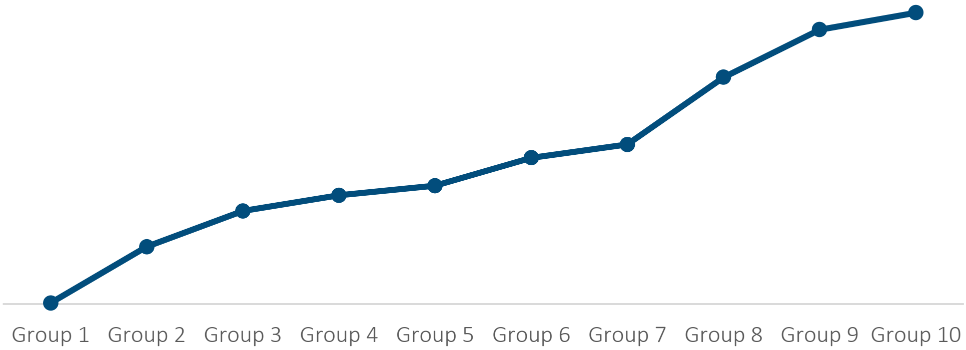
Let's be more actuarial

We are good in cutting and computing exposure and predict annual rates



What the model can bring

- During test phase, the selected policies are divided into 10 groups on average, with risk scores ranging from low to high. All these policies go through medical examination. The rates of rejection are increasing while predicted risks are increasing

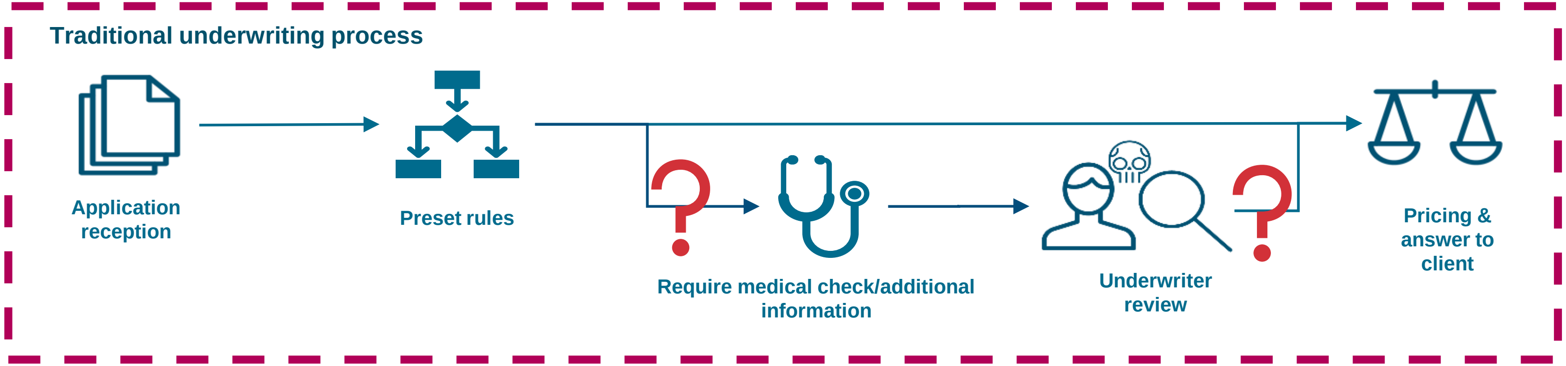


- Significant cost saving after one year on live
  - Waive the medical examination costs for good risk above X year, without increasing risk
  - With same amount of workload, the model helps to identify more bad risks

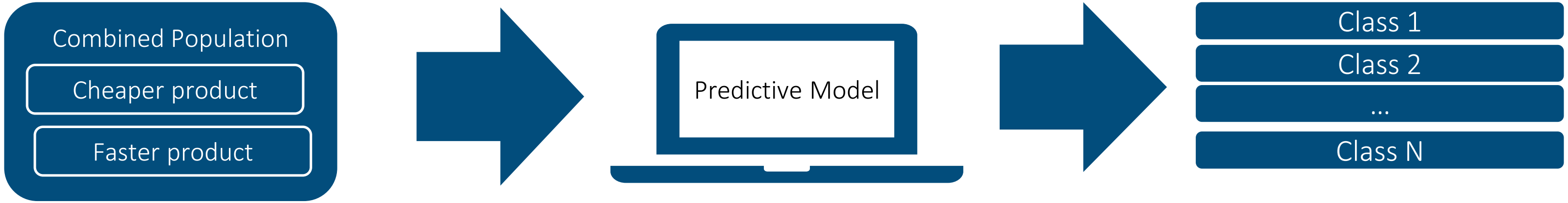
USE CASE 2: Predictive Underwriting/Pricing for Term Life

SOCIETY OF ACTUARIES®

Define the problem



- US term life market is very competitive.
- Our clients has 2 types of term products, one priced for competitiveness and one priced for fast issues. But customers always ask for cheaper rates and faster turnaround time.

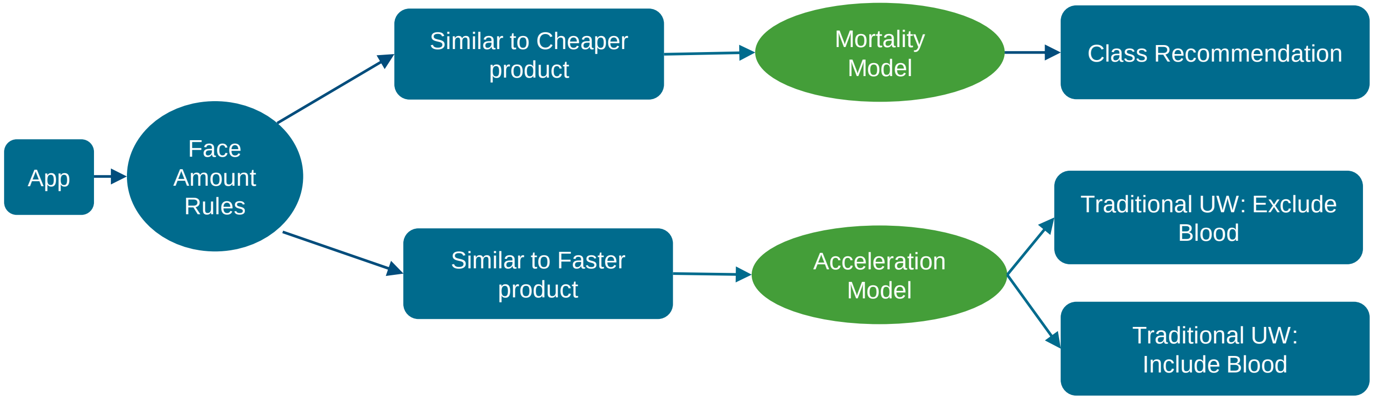


A structured answer to problem

Model each problem separately:

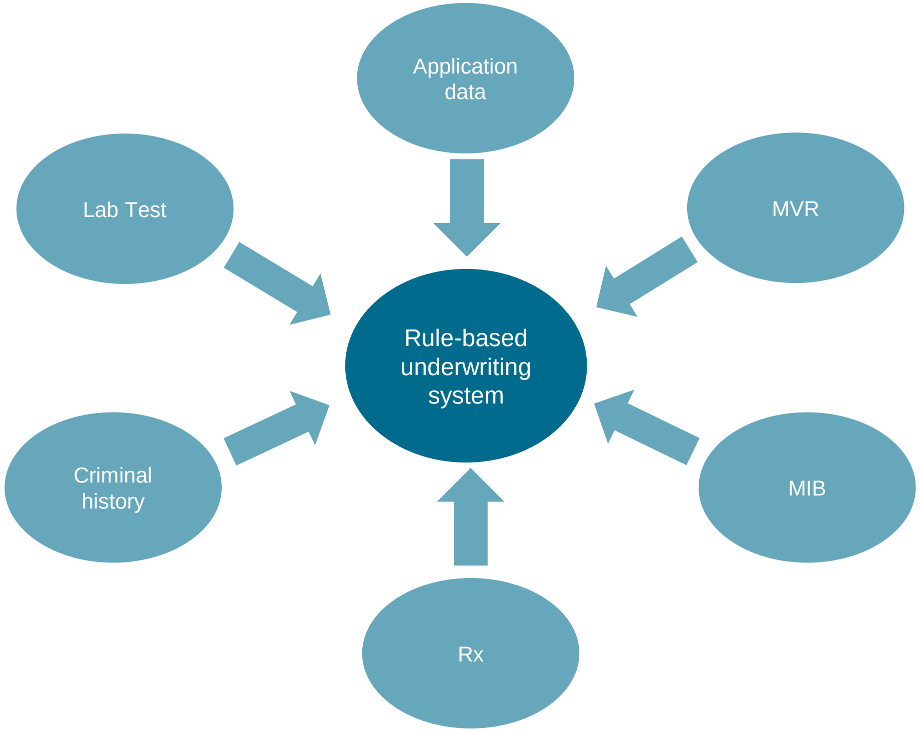
- Accelerate and reduce requirements (labs) for the cheaper product, initially waving blood as often as possible.
- Increase risk segmentation by using a mortality based predictive model for the faster product.

Create an ensemble solution:



Potential data source

- In US, common third-party data sources are available for life insurers

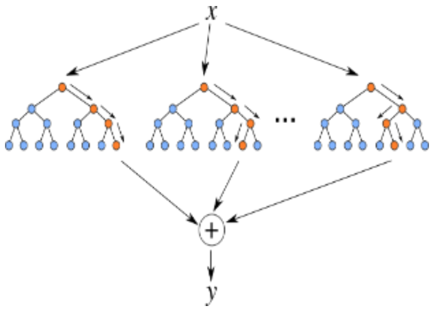


How time should be considered: A more solid approach

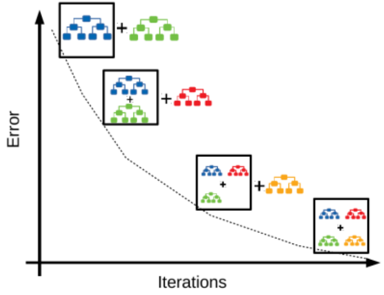
- Survival analysis: "is a branch of statistics for analyzing the expected duration of time until one or more events happen, such as death in biological organisms and failure in mechanical systems" ---From Wikipedia
- Recall Cox model:

$$\lambda(t, X_1, X_2, \ldots, X_p) = \lambda_0(t)\exp\left(\sum_{j=1}^{p} \beta_j X_j\right)$$

- Combine cox assumption with machine learning technique



Survival random forest

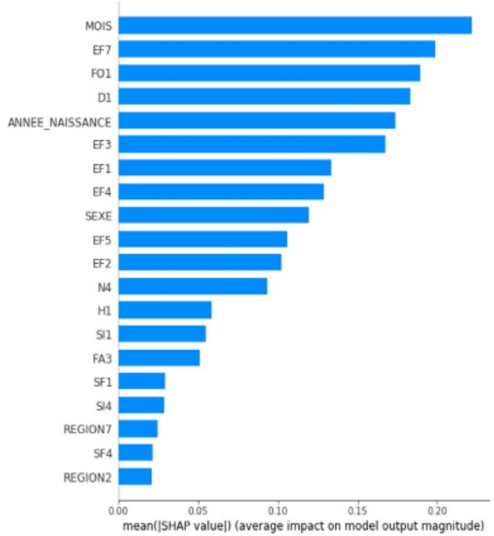


Survival XGboost

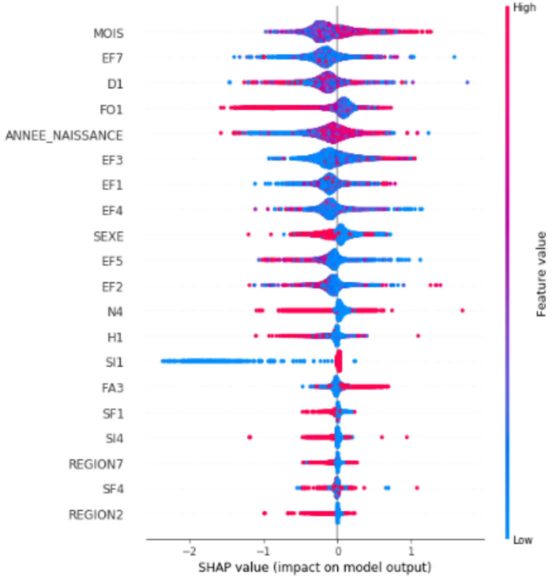
How to understand model

- A powerful interpretation tool: SHAP

e.g. Tree based model, complete Feature Importance by Interpretability tools



e.g. SHAP



SOCIETY OF ACTUARIES®