



BNP Paribas Cardif Claims Management

January 18, 2017

OVERVIEW

Problem Statement : Claims management requires different levels of check before a claim like Death, Total or Permanent Disability, Hospitalization etc. can be approved and a payment can be made.

Objective:

- To predict the category of a claim based on features available early in the process
- To help BNP Paribas Cardif accelerate its claims process
- To help BNP Paribas Cardif provide a better service to its customers.



Dataset :

- An anonymized database with two categories of claims:
 - Claims for which approval could be accelerated leading to faster payments
 - Claims for which additional information is required before approval



BUSINESS VALUE



As a global specialist in personal insurance, BNP Paribas Cardif serves 90 million clients in 36 countries across Europe, Asia and Latin America.



When faced with unexpected events, customers expect their insurer to support them as soon as possible.



With the new practices and behaviors generated by the digital economy, this process needs adaptation. Thanks to data science, the new needs and expectations of customers can now be met.



The results of the predictive model will help BNP Paribas Cardif to decide whether the claim is suitable for accelerated approval or not.

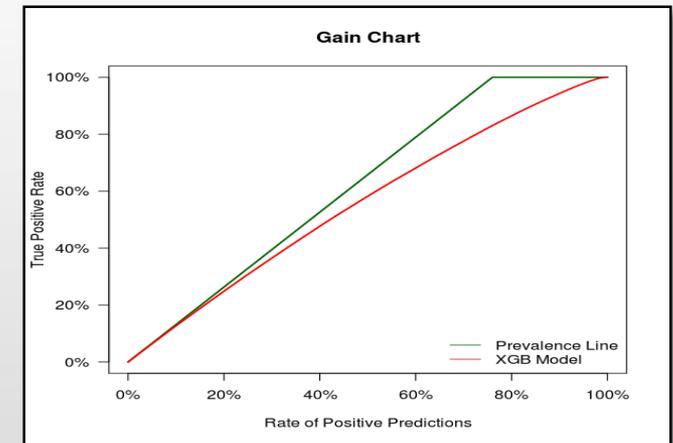
MODELING AND RESULTS

Modeling :

- The problem is a binary classification one where 1 indicates claims suitable for an accelerated approval, 0 indicates claims for which additional information is required before approval.
- XGBoost is used for training the model, of the xgboost package.
- Train data consists of 114321 clients and test consists of 114393 clients, with 131 features.

Results :

- The evaluation metric for this competition is Log Loss.
- The Log Loss with this method is 0.47.
- Log Loss quantifies the accuracy of a classifier by penalizing false classifications.
- Minimizing the Log Loss is basically equivalent to maximizing the accuracy of the classifier.



Accuracy	Sensitivity	Specificity	PPV	NPV	Kappa	Prevalence
78.02	96.02	20.06	79.32	62.4	21.2	76.11