Machine Learning For Profit Based Investing Predictive Modelling (2024 P3A)

Koen van Esterik Steven Bontius

HAN - Master Applied Data Science

7th of July 2025

Your Role in Today's Meeting

- Your Position: You are part of the Data Science leadership team of our investment company.
- The Challenge: Our telemarketing division is currently operating at a loss.
- The Proposal: We will present our research findings and a strategic plan for a turnaround.
- Your Objective: Evaluate our proposal and decide whether to approve this project.

Dear stakeholders - We are very excited to have you here today as we would like to present the findings of our research.

Problem Space

- Telemarketing division operating at a loss
- Increase the conversion rate of telemarketing calls
- Predictive modelling to select prospects
- Maximize profit from telemarketing calls

Solution Space

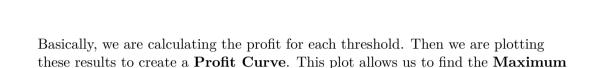
- Historical data
- Binary classification task
- Telemarketing prospect selection procedure

Introducing the Maximum Profit metric to evaluate the performance of predictive models as well as the profitability of the telemarketing division.

$$\vec{y_{pred}} = \sum_{i=1}^{thresholds} \begin{cases} 1 & \text{if } \vec{y_{probs}} \ge threshold_i \\ 0 & \text{otherwise} \end{cases}$$

$$ec{tps}, ec{fns}, ec{fps}, ec{tns} = \sum_{i=1}^{thre ec{sholds}} ext{confusion_matrix}(y_{ec{true}}, y_{ec{pred}})$$

 $total\ profit = \max{(profit\ per\ subscription*t\vec{ps} - cost\ per\ call*f\vec{ps})}$



Profit.

Methodology

- Data Preprocessing
- Exploratory Data Analysis
- Splitting Data
- Cross Validating Models
- Selecting Model
- Evaluating Model
- Conclusion

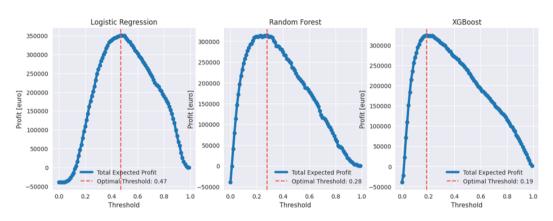


Figure 1: Model Shortlist

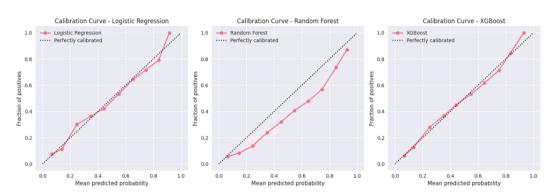


Figure 2: Model Calibration

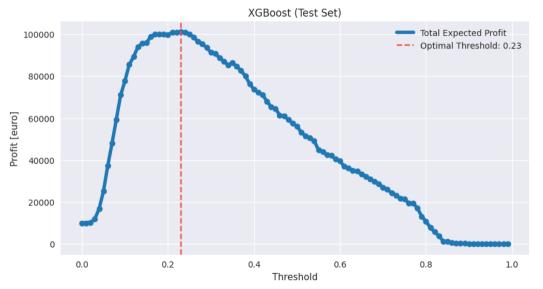


Figure 3: Model Selection

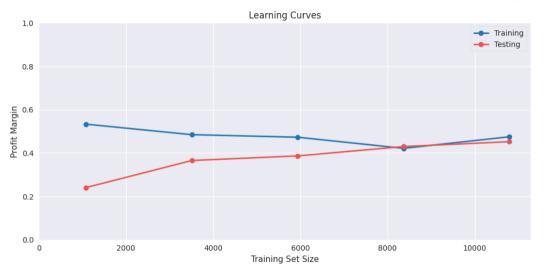


Figure 4: Learning Curves

Conclusion

Comparison of Profit with and without Predictive Modelling:

Procedure	Profit
Call All Prospects	10,000
Call Preselected Propspects	101,100

Conclusion

"Gas op die lolly?"

Remarks

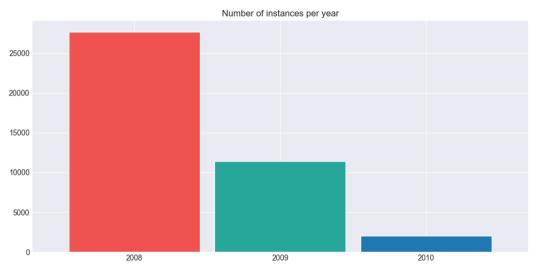


Figure 5: Instances per Year $\,$

Remarks

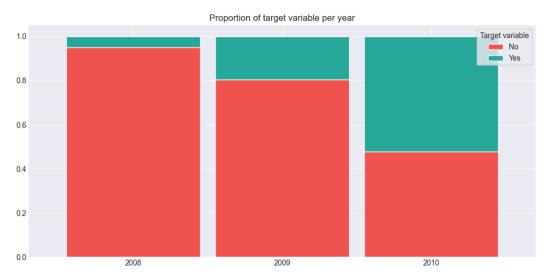


Figure 6: Proportion of Target Variable per Year

Remarks

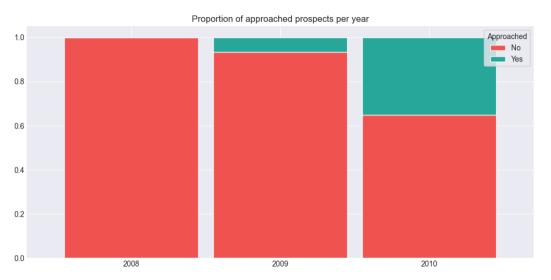


Figure 7: Proportion of Approached Prospects per Year

