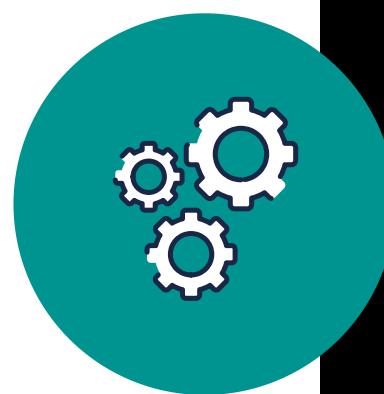


# Estimating Clients' Needs



# Contents



## Exploration and preprocessing

What data do we have ? How do we use it ? Only one way to find out : exploration and interpretation.

## Modeling

Use of Machine learning algorithm to predict the needs of our future customers.

## Next Best Action

Use these predictions to get closer to customers

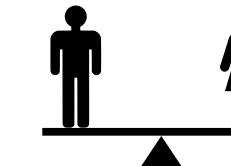
# Our data :

2 targets : Accumulation and  
Income Investment

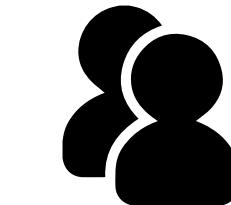
5000 customers



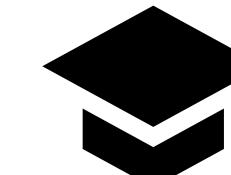
Age



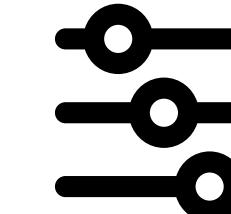
Gender



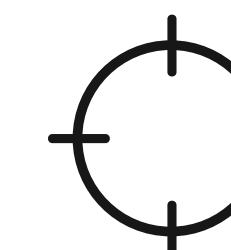
Family Members



Financial Education



Risk Propensity



Income



Wealth

# Preprocessing



Garbage in garbage out !

## Adaptation

Income and wealth

Application of a log transformation  
on these data for a better  
performance of the algorithms

## Modification

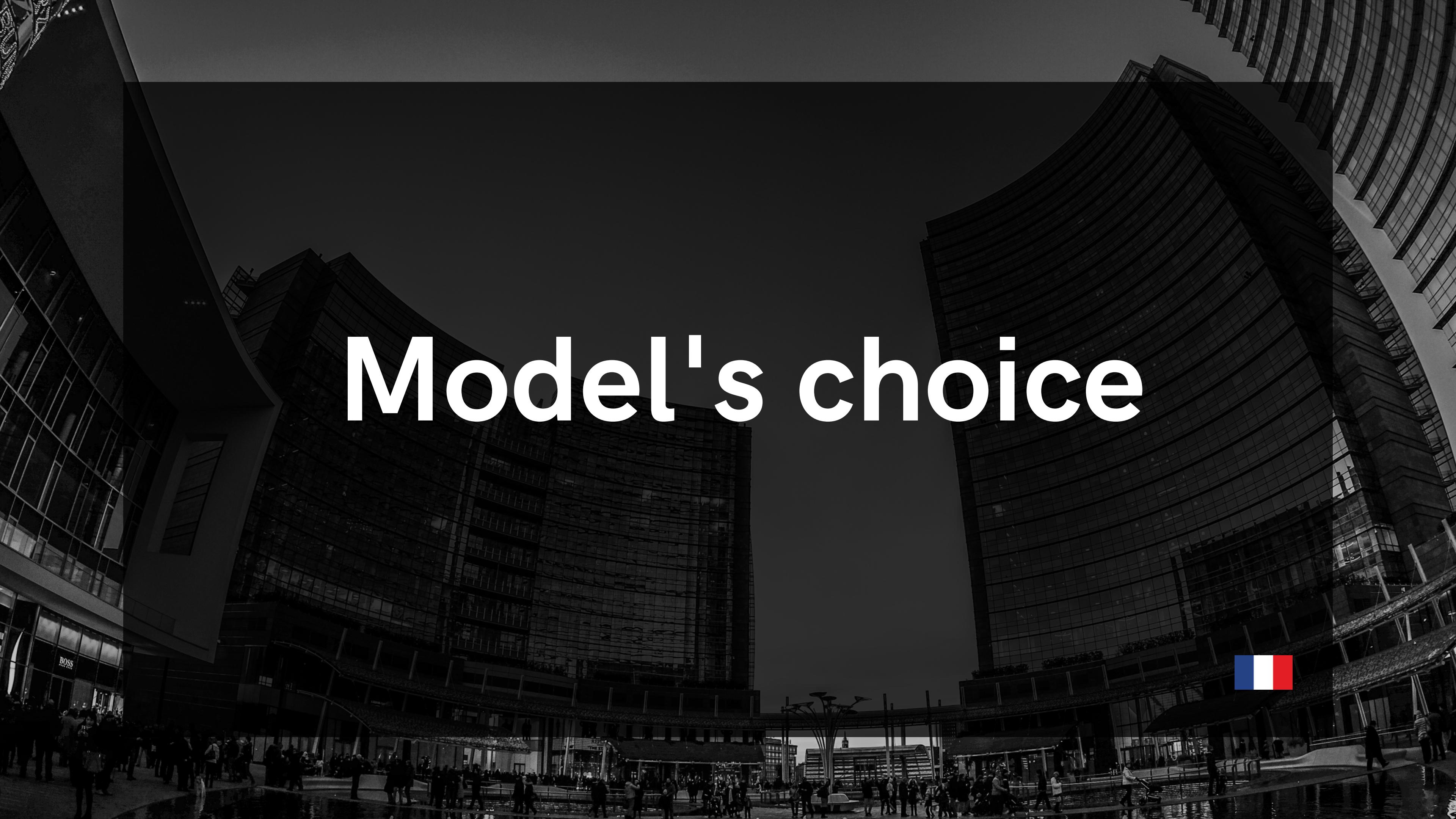
Education and risk

Implementation of a ratio between  
these features to factorize  
information due to the correlation

## Supression

Gender & Family members

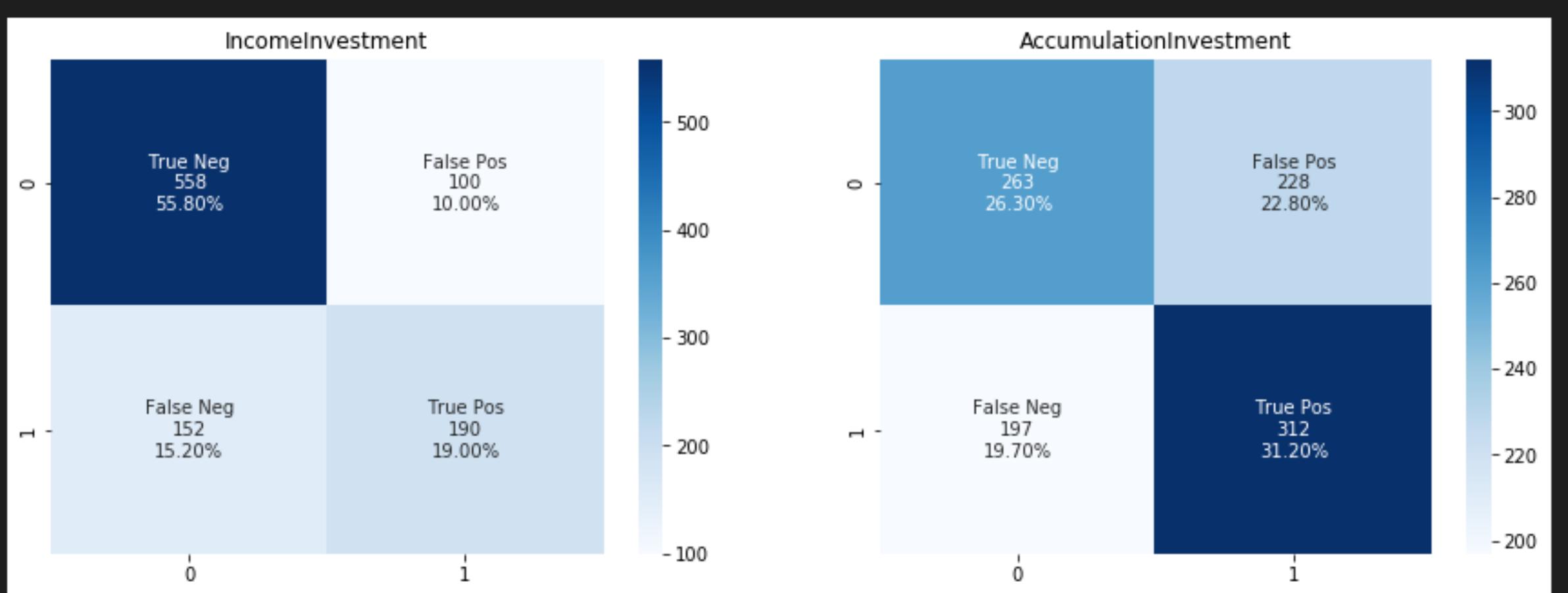
Removal of these features as no  
addition of information on a  
business point of view



Model's choice



# First Results



Low results for simplest models

Measures for IncomeInvestment:

Sensitivity = 55.56% ; Specificity = 84.80% ; Precision = 65.52% ; F1 score = 60.13% ; Accuracy = 74.80%

Measures for AccumulationInvestment

Sensitivity = 61.30% ; Specificity = 53.56% ; Precision = 57.78% ; F1 score = 59.49% ; Accuracy = 57.50%

# Key metrics

## Accuracy

Keep global performances

How well a binary classification test correctly identifies or excludes a condition : the proportion of correct predictions

## Recall

Avoiding waiting customers

It is the fraction of the relevant items that are successfully retrieved.

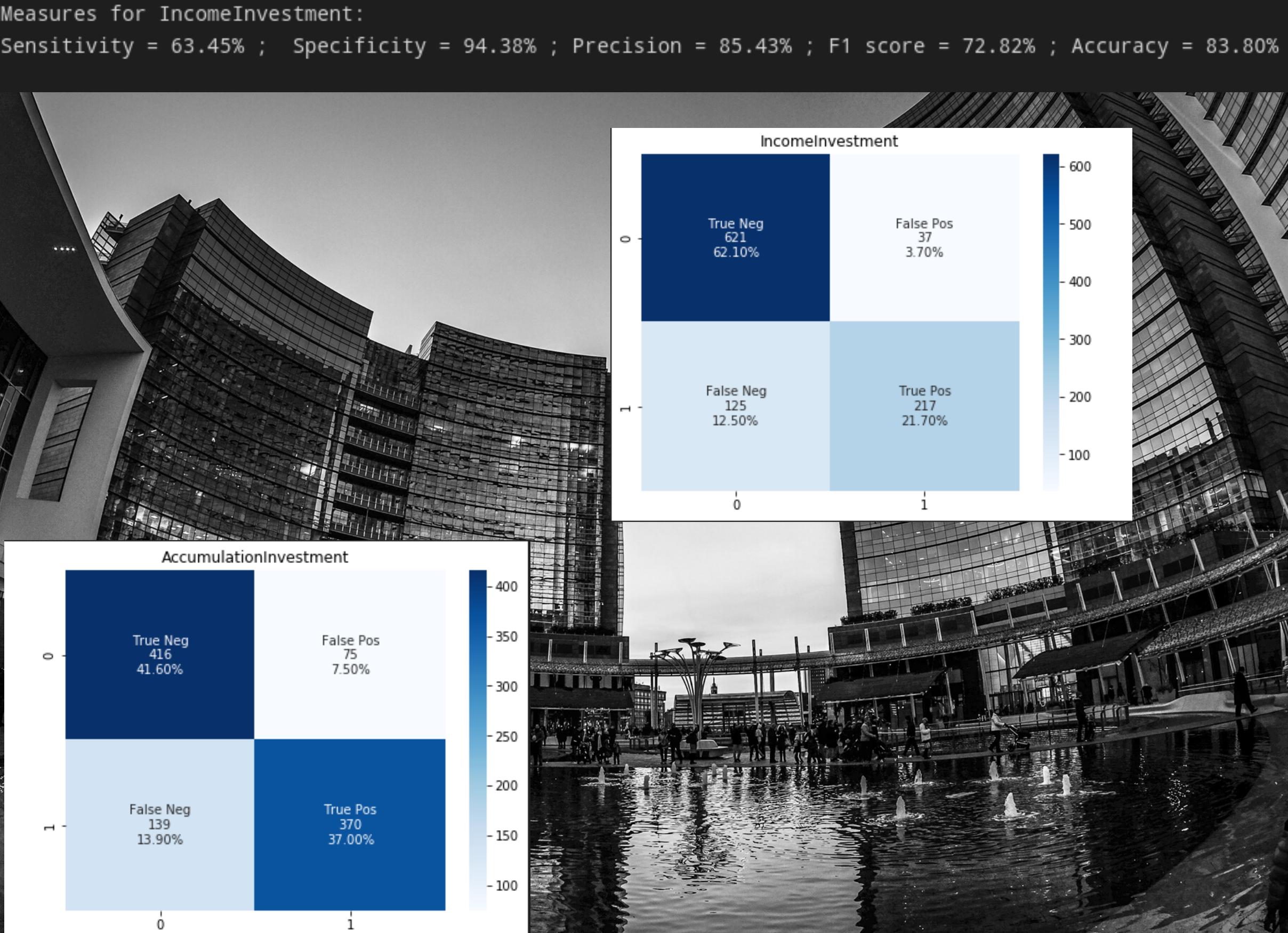
## Z-Metric

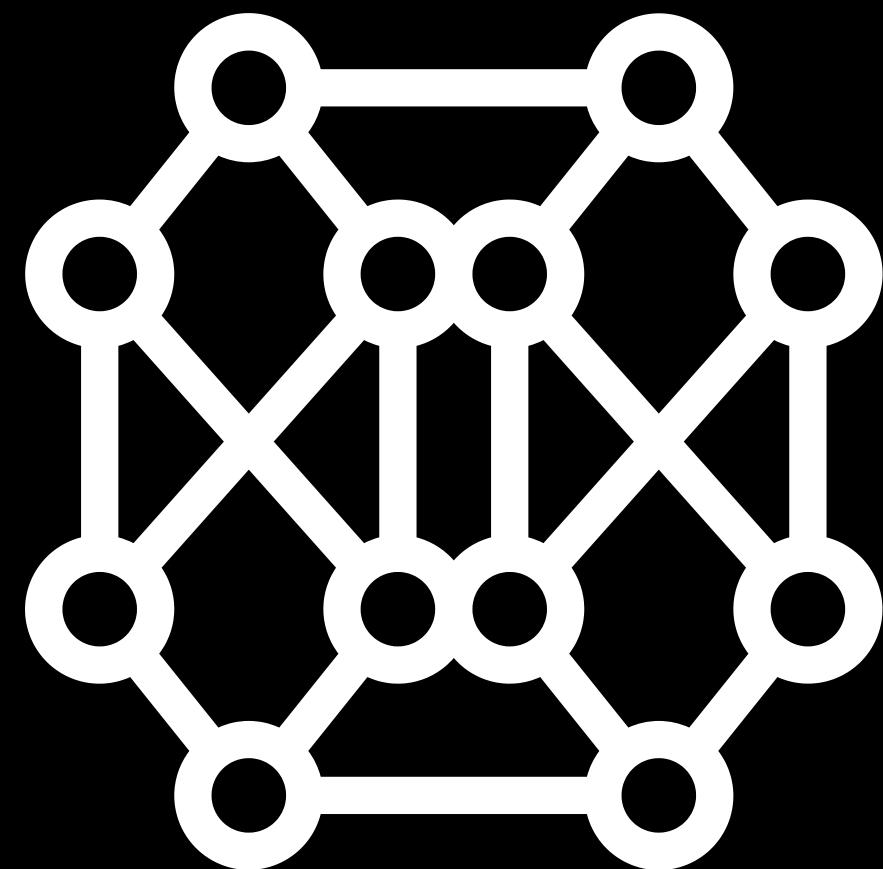
Business metric

The idea of this metric is to evaluate a model thanks to its misclassified prediction and encourage to have false positive than false negative

# Calibration

Based on these different metrics, we have process models and select which one to use for our targets

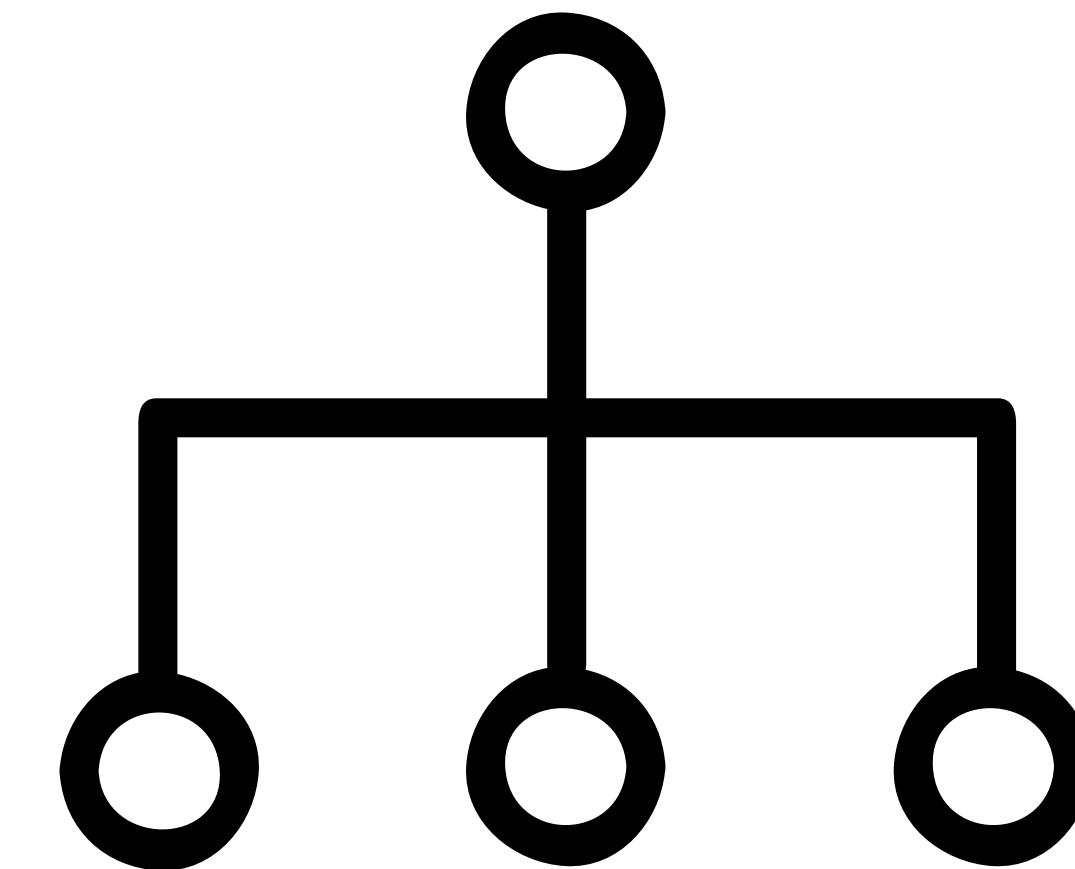




# Neural Network

Accumulation Investment

Optimization and ensemble of ANN



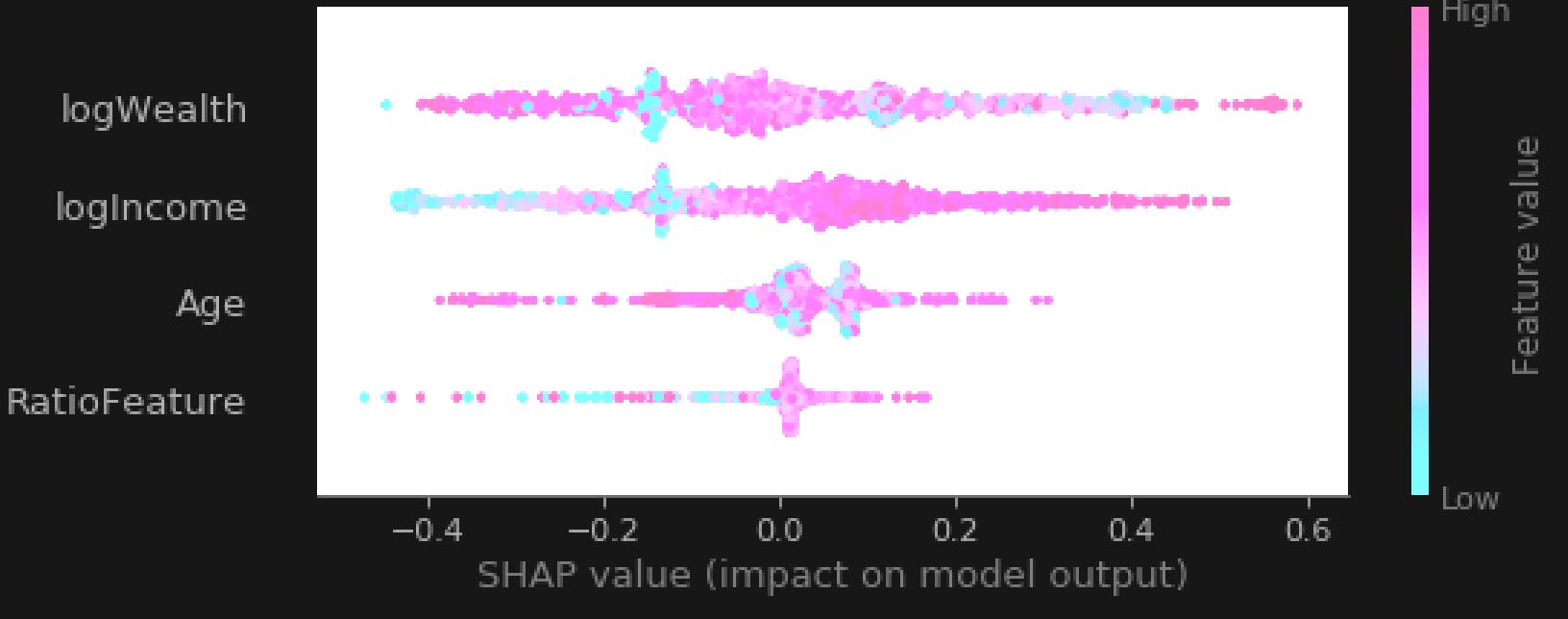
# Random Forest

Income Investment

Optimization and interpretability

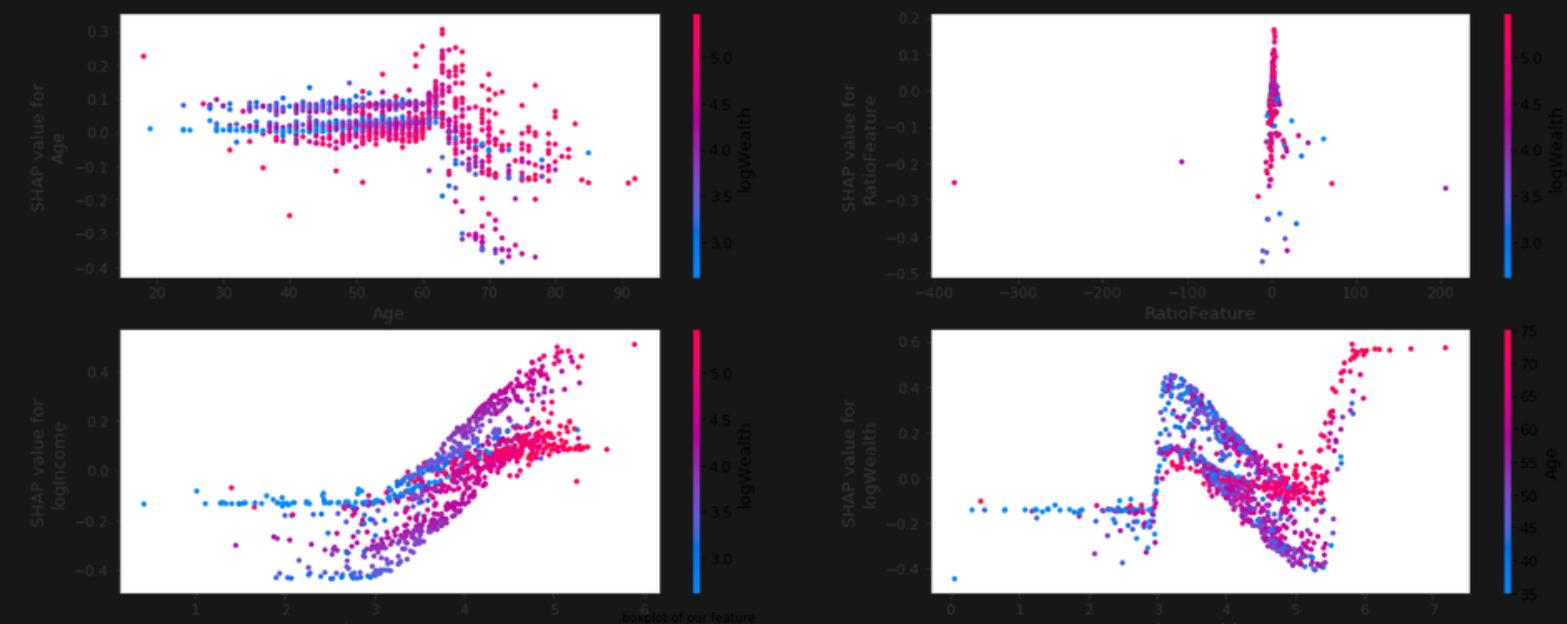
# Interpretability

GLOBAL

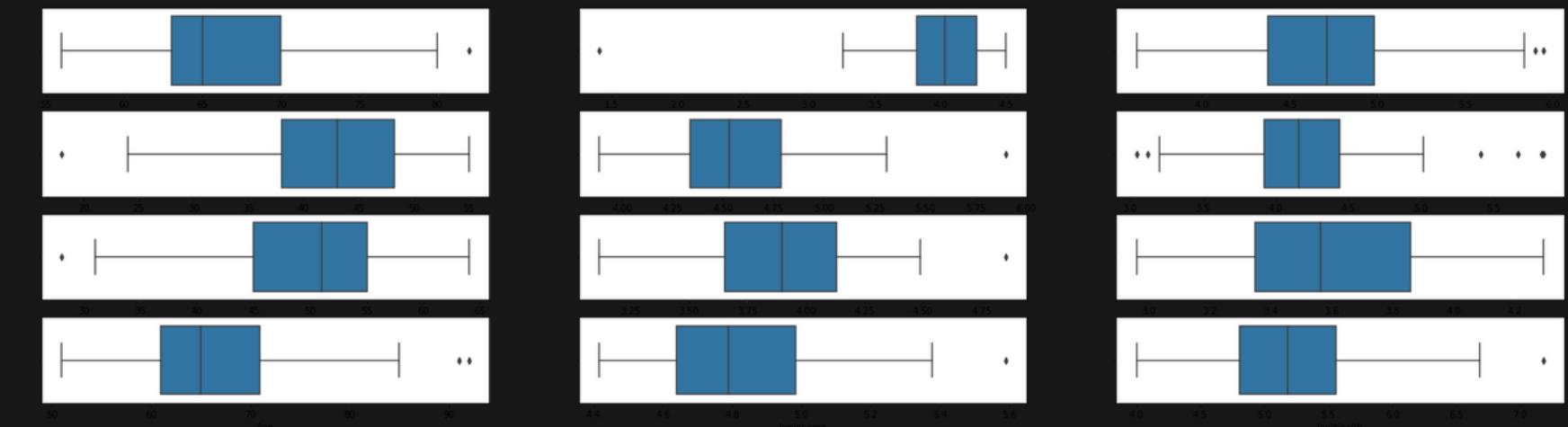


Use of features importance  
and shap values

LOCAL



Clustering on predictions  
and boxplot on clusters



# Next Best Action



# Risk scale

All our business plan stand on a risk scale

- A first risk number attributed to each client that describes his probability of default
- Four different risks that describe the probability of default of the four product provided by the company



# Business choice

**Between the four products that can be advised to the client we choosed a specific strategie**

- **Always advise the closest "less risky" if it exists**
- **Choose between the closest "more risky" and the second closest "less risky" following distance criteria**



# Grazie mille!

Contattaci per qualsiasi domanda.

Saâd Elhaimar  
Julien Oualid  
Emmalito Delar  
Julien Dritsas

