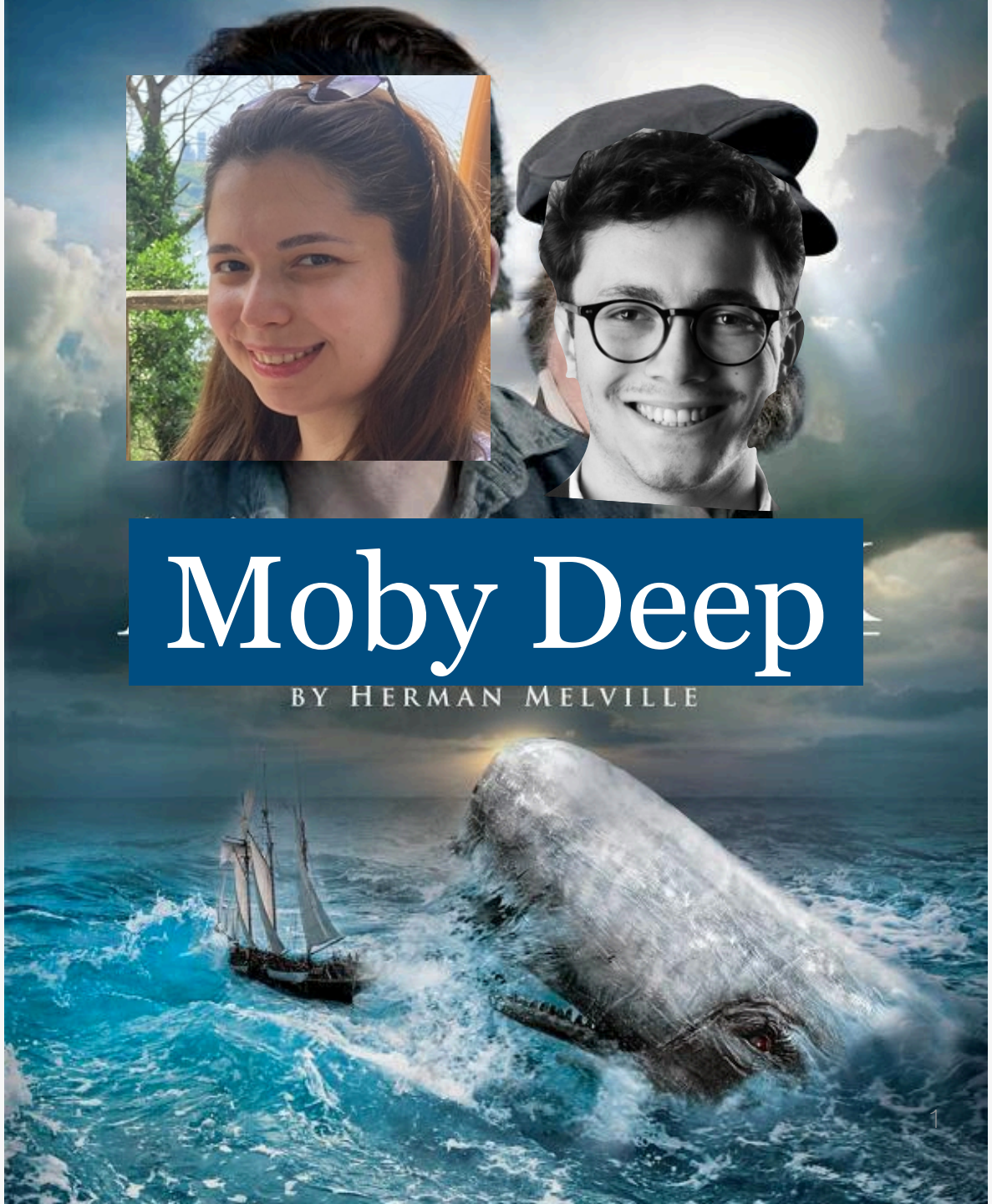

Moby Deep
BY HERMAN MELVILLE

Classification | Bank project

How to know which client will say **“Yes”** or **“No”**
to our offers ?

Table of Contents

Tableau Exploration

01

Offer Accepted | Dashboard

https://github.com/thomasmaechler/Case-Study-Classification

Offer Accepted / Credit card rating

Offer Accepted	Low	Medium	High
No	31.36%	33.58%	35.06%
Yes	61.88%	26.49%	11.63%

Offer accepted / Mailer Type

Offer Accepted	Letter	Postcard
No	50.38%	49.62%
Yes	29.33%	70.67%

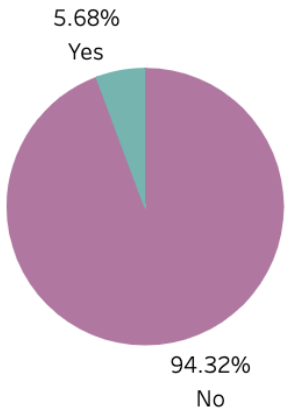
Offer accepted / Overdraft Protection

Offer Accepted	No	Yes
No	85.10%	14.90%
Yes	85.53%	14.47%

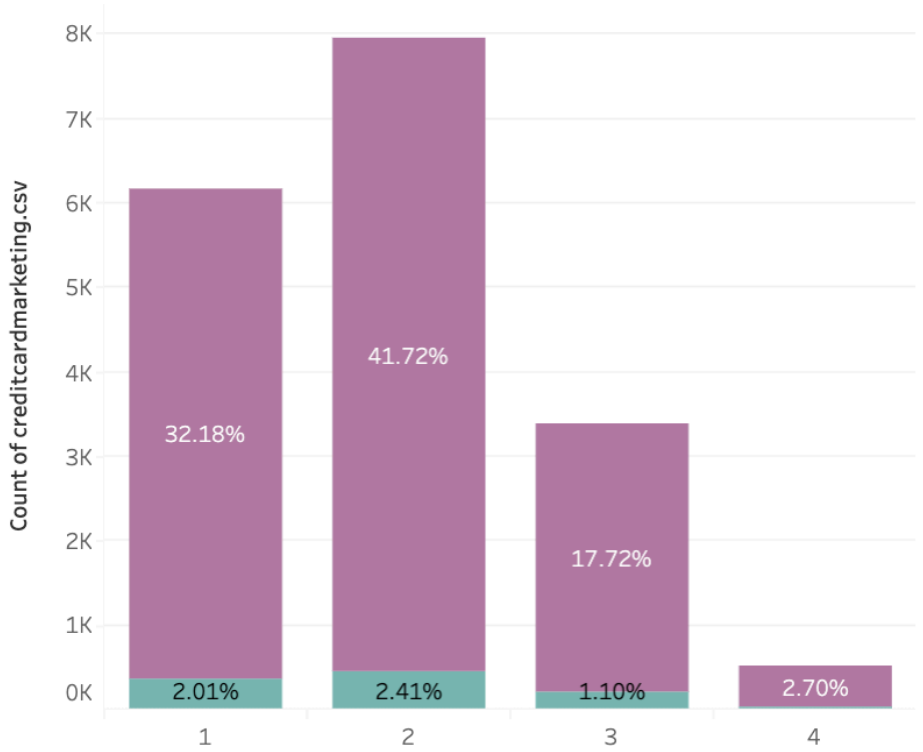
Offer accepted / Reward

Offer Accepted	Air Miles	Cash Back	Points
No	32.96%	34.12%	32.92%
Yes	45.45%	20.14%	34.41%

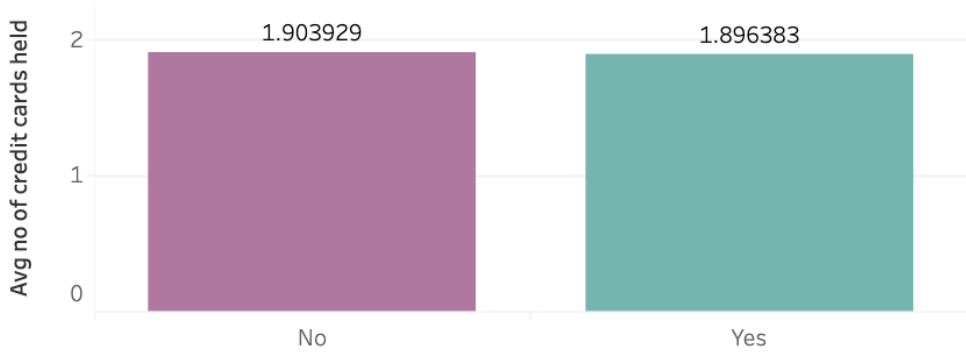
Percentage of people saying "Yes" or "No"



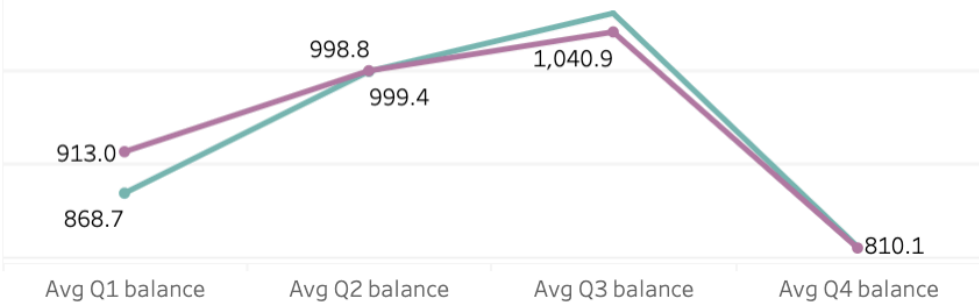
% of "Yes/No" over n° of credit cards held



Average number of credit cards held



Average Balances per Quarter



Python Discoveries

02

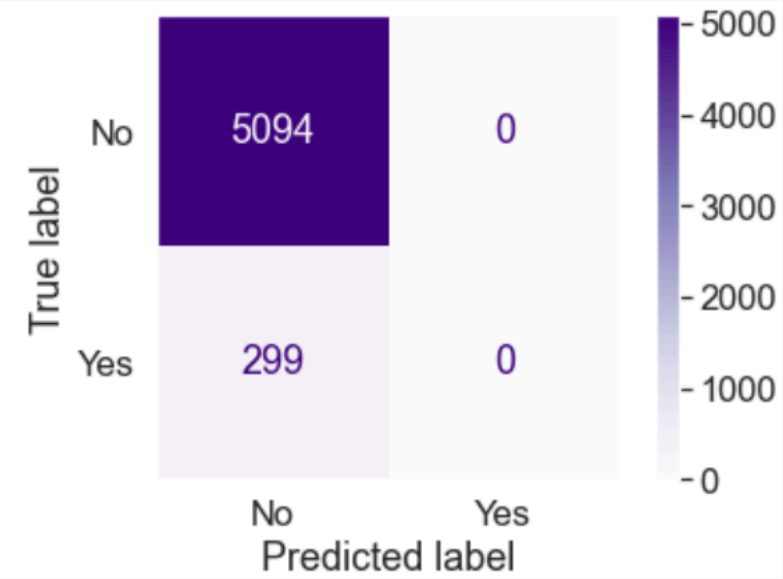
Problems & data

Steps

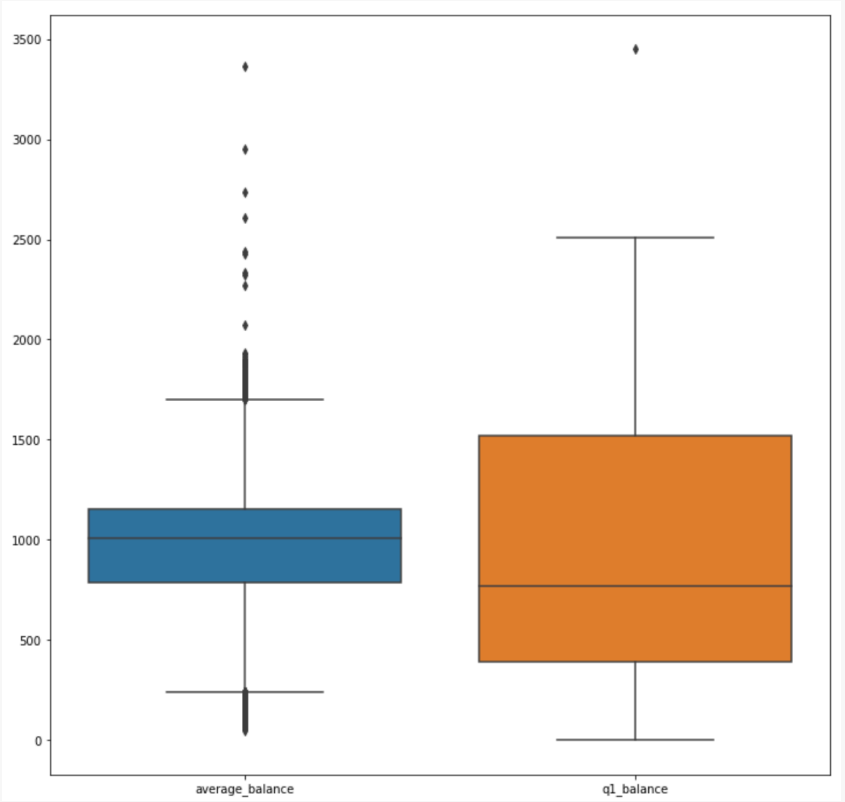
- EDA
- Correlation matrix
- 2 Jupyter Notebooks : With/Without outliers
- Dealing with the features to take or not
- Choosing the right Scaling method STD / Normalizer / MinMax
- Choosing the right Sampling method
- Dealing with the Unbalanced data

Logistic regression

Confusion Matrix



Plot figures



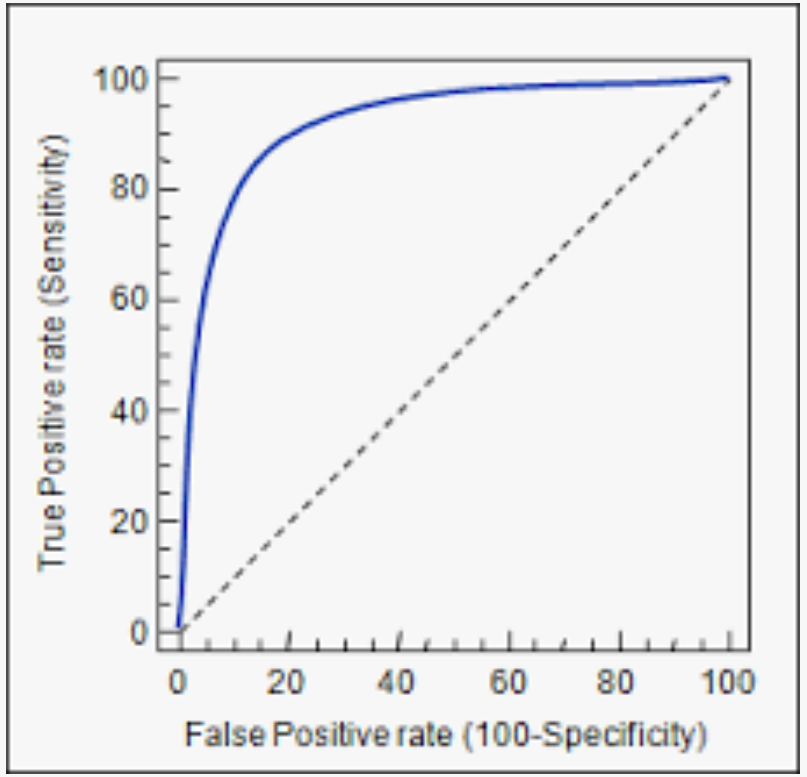
Our mistakes

1 - Not understanding well

- We used a different sampling methods : “BalancedBaggingClassifer”

2 - Overfitting

At the end we understood that the model was overfitting



Our models

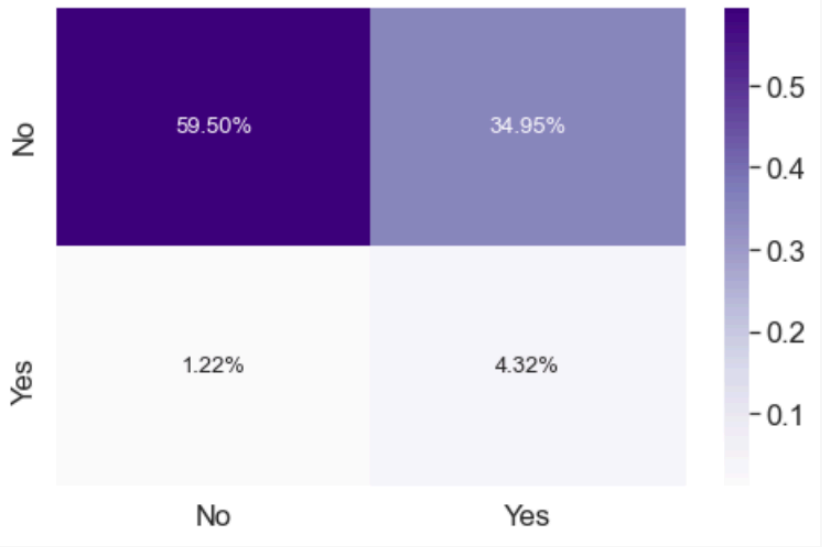
	OUR BEST MODEL	Model 1	Model 2	Model 3
Columns removed	no_of_bank_accounts_open - q2_balance -q3_balance - q4_balance - no_of_homes_owned- no_of_credit_cards_held - household_size			
Scalers	Normalizer	Normalizer	Normalizer	MinMax
Outliers	With	With	With	Without
Sampling methods	SMOTE - TOMEK LINKS	KNN	SMOTE	SMOTE
Results	Accuracy Score = 0.64 AUC Score = 0.77 F1 Score = 0.84	Accuracy Score = 0.94 AUC Score = 0.60 F1 Score = 0.52	Accuracy Score = 0.67 AUC Score = 0.77 F1 Score = 0.76	Accuracy Score = 0.67 AUC Score = 0.76 F1 Score = 0.79

Our Best Model

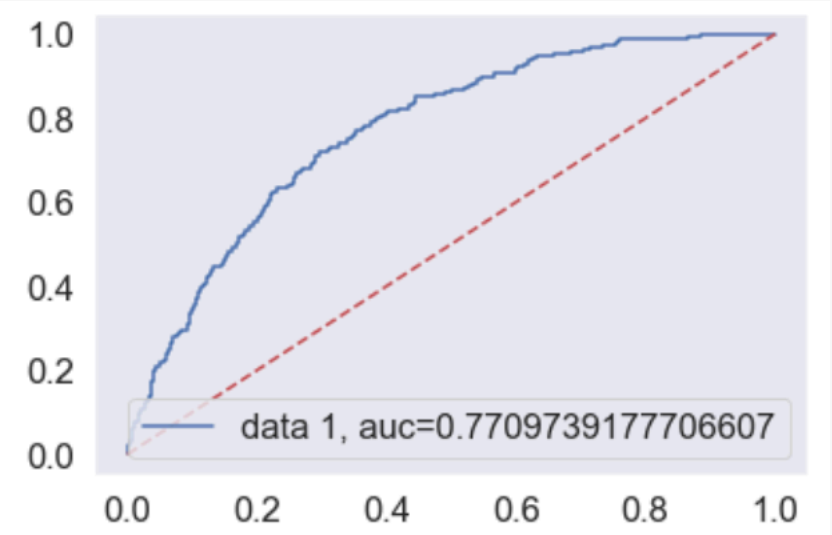
03

Our best model

Confusion Matrix



ROC Curve



Results

Summary of the results of this model:

- accuracy score = 0.64
- auc score = 0.77
- f1 scores are around 0.84.

Do you have any questions ?

Thanks