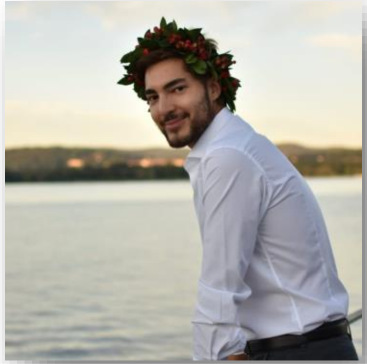
HACKATHON 2021 Monabanq

TEAM APK

27/05/2021

Our Team Members

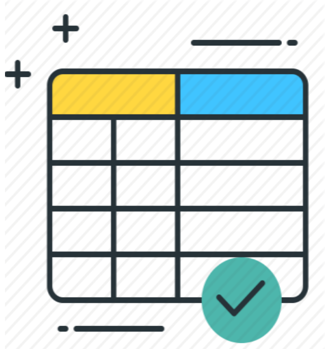


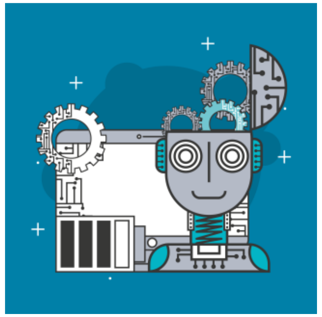
Alessandro Varnelli

Prineet Kaur

Kiran TL

Table of Contents





Project Goal

- Problem Statement
- Business Target
- Desired Outcome

Timeline

- Timeframe Considered
- Dependent Variable
- Independent Variables

Basetable

- Preprocessing
- Properties
- Final Design & Logic

Modelling

- ML Setup
- Performance
- Finalizing

Conclusions

- Final Dashboard
- Customer Journeys
- Business Takeaways

IÉSEG SCHOOL OF MANAGEMENT

Business Goal Definition

Understanding Customer Web Journey on Monabanq and Provide Useful Insights to help optimize their Marketing Efforts in order to Boost their Online Credit Subscriptions.

Understand the Navigation Patterns of Customers

Identify Customers that need to be pushed to Convert

Contact the Targeted Customers to Buy Subscription

IÉSEG SCHOOL OF MANAGEMENT

Project Timeline

The Target Period is represented as One Week.

To Predict conversion during the Target Week, the available customer information and the past navigation behavior was used. This can be graphically represented as:

Historical Behaviour for Target Week – 4 for Customer X	Historical Behaviour for Target Week – 3 for Customer X	Historical Behaviour for Target Week – 2 for Customer X	Historical Behaviour for Target Week – 1 for Customer X	Did Customer X convert in Target Week?

This was repeated for all customers (for each week) that we had in our dataset and each one of this repetition was captured as an observation in the Final Basetable.

IÉSEG SCHOOL OF MANAGEMENT

Dependent Variable

GOAL: Did Customer x did Purchase of Credit Limit in Week z?

HOW: Analysing the Behaviour of Customer x in the past 4 weeks

ACTION: Target Variable obtained by looking at the Credit Subscription Date

Distribution of the Target Variable values –

97% Credit Not Purchased

3% Credit Purchased

* The Target column has values of 0's & 1's where -
  0 = Credit Limit not purchased,
  1 = Credit Limit purchased

IÉSEG SCHOOL OF MANAGEMENT

Independent Features in the Basetable

Profile Related Features:

- Age group
- Risk index
- Access type
- Customer segment

Navigation Related Features for each of the last 4 weeks:

- Product related pages seen
- Total number of pages seen
- Number of credit-related pages seen

Structural Features (not used in machine learning):

- Client ID
- Current target week analysed

Total number of features: **136**

IÉSEG SCHOOL OF MANAGEMENT

Preprocessing the Raw Data

- Cleaned the data to simplify some features and remove redundancies (agency codes etc.)
- Sorted page visits by Client ID and the timestamp of the visit
- Grouped page visits by Client ID and Week of visit
- For each week of navigation saved whether customers that visited the website converted or not
- Recursive data extraction for past weeks aggregated data

* If a Customer didn't have any page visits for a specific week, we don't have an observation for it in the Basetable.

Final Basetable

From all the customer attributes being captured on Monabanq website, our team brainstormed, processed and considered the most relevant ones for further analysis to predict their credit purchase intent.

Enriched

Most Relevant Features extracted being used

Simple

Easy to Interpret (all redundancies removed)

Complete

For Customer Analysis and Insights

Clean

Filtering done in line with Project Timeline

IÉSEG SCHOOL OF MANAGEMENT

Final Basetable Snapshot

Client_ID	Target	Week	Access_Type	Account_Type	Is_Customer	Risk_Index	Activity_Level	Amt_Awarded	Amt_Requested	...	Comment financer ses projets déco_Week_T-2	Comment financer ses projets déco_Week_T-1	Tout savoir MaPrimeRenov' : vos travaux financés_Week_T-4	Tout savoir MaPrimeRenov' : vos travaux financés_Week_T-3	Tout savoir MaPrimeRenov' : vos travaux financés_Week_T-2	Tout savoir MaPrimeRenov' : vos travaux financés_Week_T-1	Souscrivez à un crédit travaux| Monabanq_Week_T-4	Souscrivez à un crédit travaux| Monabanq_Week_T-3	Souscrivez à un crédit travaux| Monabanq_Week_T-2	Souscrivez à un crédit travaux| Monabanq_Week_T-1
1	0	44	CMDFORF	PERSONNE_PHYSIQUE	True	D+	GESTION	missing	missing	...	0	0	0	0	0	0	0	0	0	0
2	0	44	CMDFORF	PERSONNE_PHYSIQUE	True	A-	GESTION	missing	missing	...	0	0	0	0	0	0	0	0	0	0
3	0	44	CMDFORF	PERSONNE_PHYSIQUE	True	D-	GESTION	missing	missing	...	0	0	0	0	0	0	0	0	0	0
4	0	44	CMDFORF	PERSONNE_PHYSIQUE	True	C+	GESTION	missing	missing	...	0	0	0	0	0	0	0	0	0	0
6	0	44	CMDFORF	PERSONNE_PHYSIQUE	True	B-	GESTION	missing	missing	...	0	0	0	0	0	0	0	0	0	0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
63360	0	49	CMDFORF	PERSONNE_PHYSIQUE	True	A-	GESTION	missing	missing	...	0	0	0	0	0	0	0	0	0	0
63361	0	49	CMDFORF	PERSONNE_PHYSIQUE	True	A-	GESTION	missing	missing	...	0	0	0	0	0	0	0	0	0	0
63362	0	49	CMDFORF	PERSONNE_PHYSIQUE	False	missing	GESTION	missing	missing	...	0	0	0	0	0	0	0	0	0	0
63478	0	49	CMDFORF	PERSONNE_PHYSIQUE	True	E+	GESTION	missing	missing	...	0	0	0	0	0	0	0	0	0	0
63488	0	49	CMDFORF	PERSONNE_PHYSIQUE	True	B-	GESTION	missing	missing	...	0	0	0	0	0	0	0	0	0	0

Final Size for Project Analysis: **125.189 observations × 136 features**

Machine Learning Setup

- We performed a combination of **Oversampling and Undersampling** to Increase the Recall of the minority class**.**

- **Grid Search Parameters:**

  logistic = LogisticRegression(C= 100, penalty= 'l2', solver= 'newton-cg’)

  neighbors = KNeighborsClassifier(algorithm= 'kd_tree', leaf_size= 10, metric= 'manhattan', n_neighbors= 35, weights='distance’)

  randomForest = RandomForestClassifier(max_features='log2', n_estimators=600)

  neuralNet = MLPClassifier(alpha= 0.01, hidden_layer_sizes= 14, max_iter= 1800, random_state= 6, solver= 'lbfgs’)

  GradientBoost = GradientBoostingClassifier(criterion='friedman_mse', learning_rate= 0.2, loss= 'deviance', max_depth= 8, max_features= 'sqrt’, min_samples_leaf=2, min_samples_split= 0.2, n_estimators= 150, subsample= 0.95)

- Categorical variables encoding through **Ordinal Encoder:** This method allows for a quick and easy encoding by assigning a numerical value to each of the classes for each of the categorical variables.

- Feature Selection through **Pearson test:** In our model we consider just 47 columns, so it is quick to run and still interpretable.

IÉSEG SCHOOL OF MANAGEMENT

Model Performance

Performance on Test Set:

F1 score is still quite low, but we have 15% True positive in our predictions.

	logistic	neighbors	randomForest	neuralNet	GradientBoosting
AUC	0.723324	0.674394	0.723792	0.684951	0.761734
F1	0.054054	0.026538	0.037773	0.028966	0.000000
Precision:	0.030928	0.014493	0.020299	0.015217	0.000000
Recall for minority class:	0.214286	0.157143	0.271429	0.300000	0.000000

After analysing the above we decided to go ahead with "**Logistic Regression**" for Final Model

Business Interpretation:

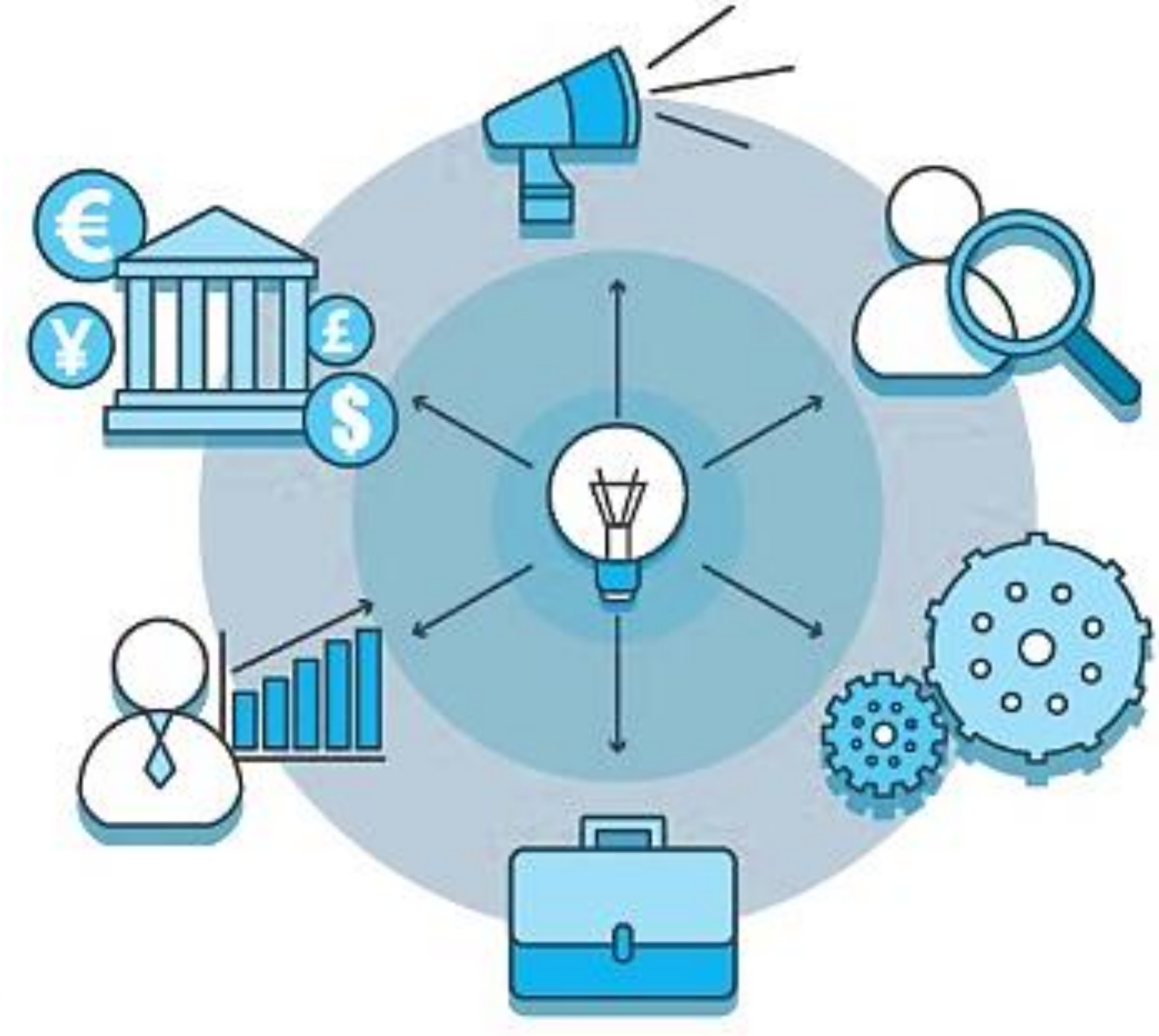
If we contact the people identified, at least the 15% will convert.

For our test set, this potentially represents ~18k* Euro of conversion.

*(6 [15% of the people found in the test set] x 3k [= most popular Amount Requested])

IÉSEG SCHOOL OF MANAGEMENT

Visualizing the Key Findings (Dashboard & Customer Journey)



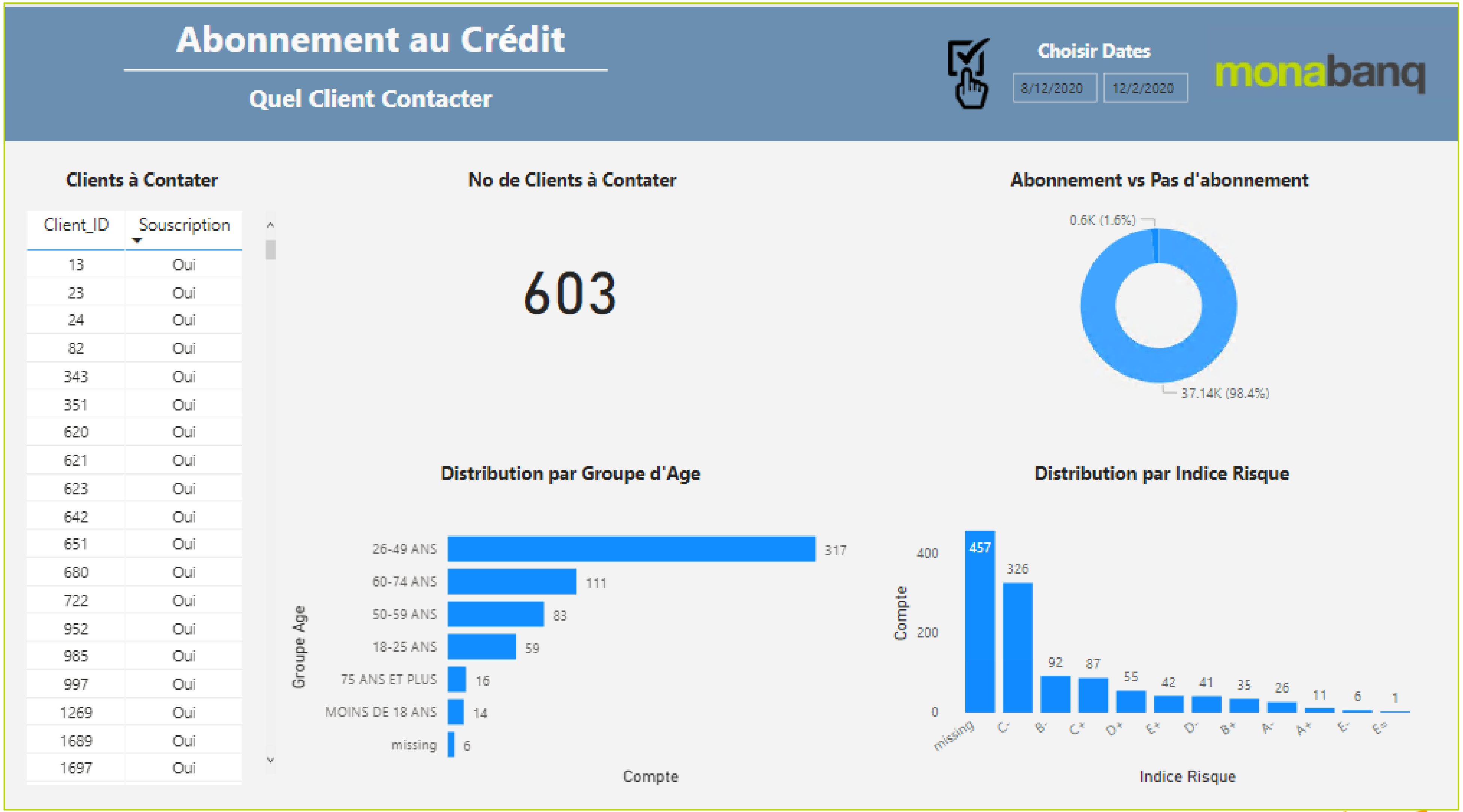
IÉSEG SCHOOL OF MANAGEMENT

Dashboard

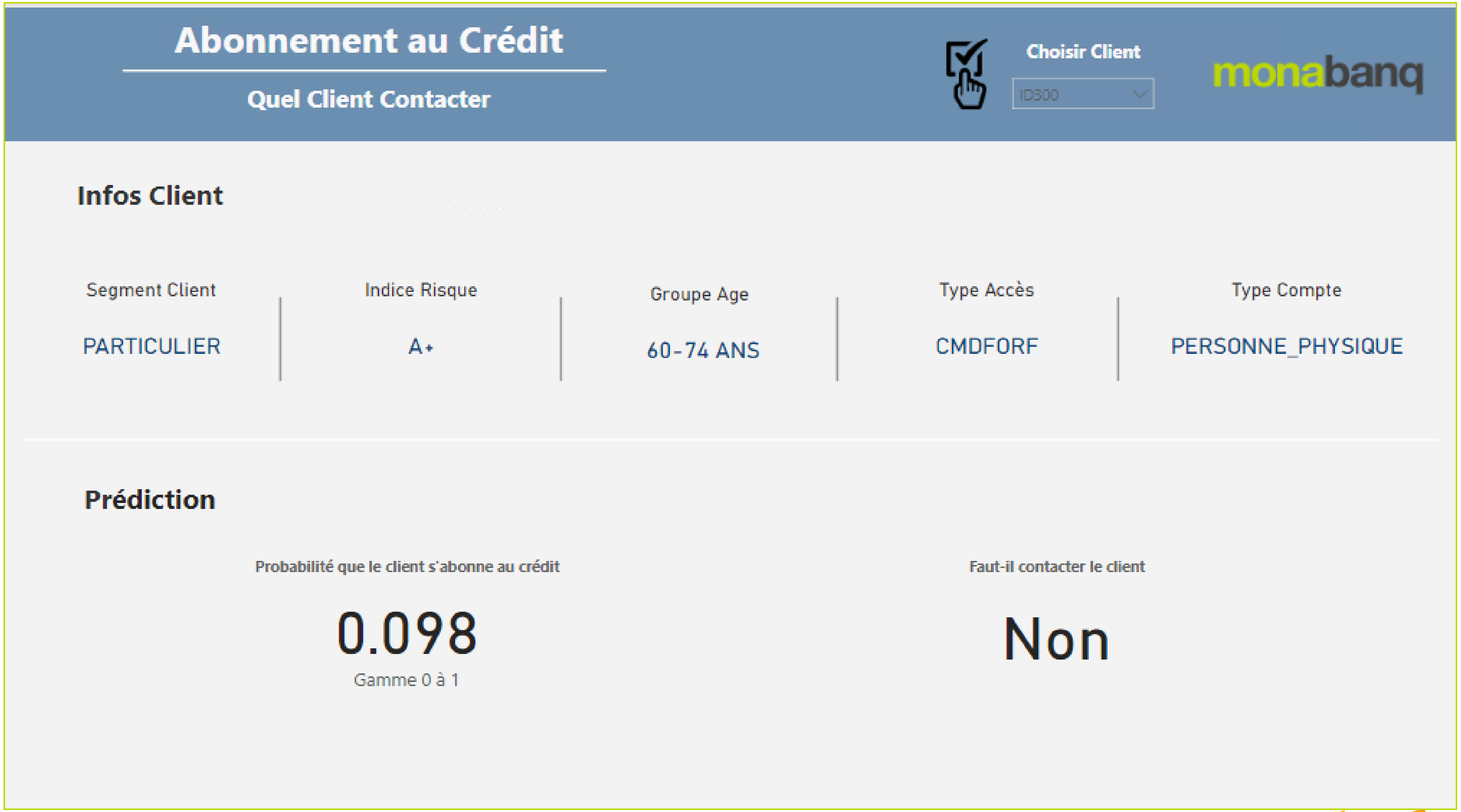
Key reasons that we decided to use Power BI –

- ❑ No Memory or Speed Constraints
- ❑ High Level of Personalization can be easily achieved
- ❑ Integrated easily with the existing application
- ❑ Simple and Secure to build Reports
- ❑ No Specialized or External Technical support required

Overall Trend (Weekly/Date wise):



Customer Specific Details:



Mapping the Customer Journey

Aims to help Monabanq better visualise interactions with their visitors, in order to let it's Marketing team plan and optimise the whole website experience throughout the touchpoints.

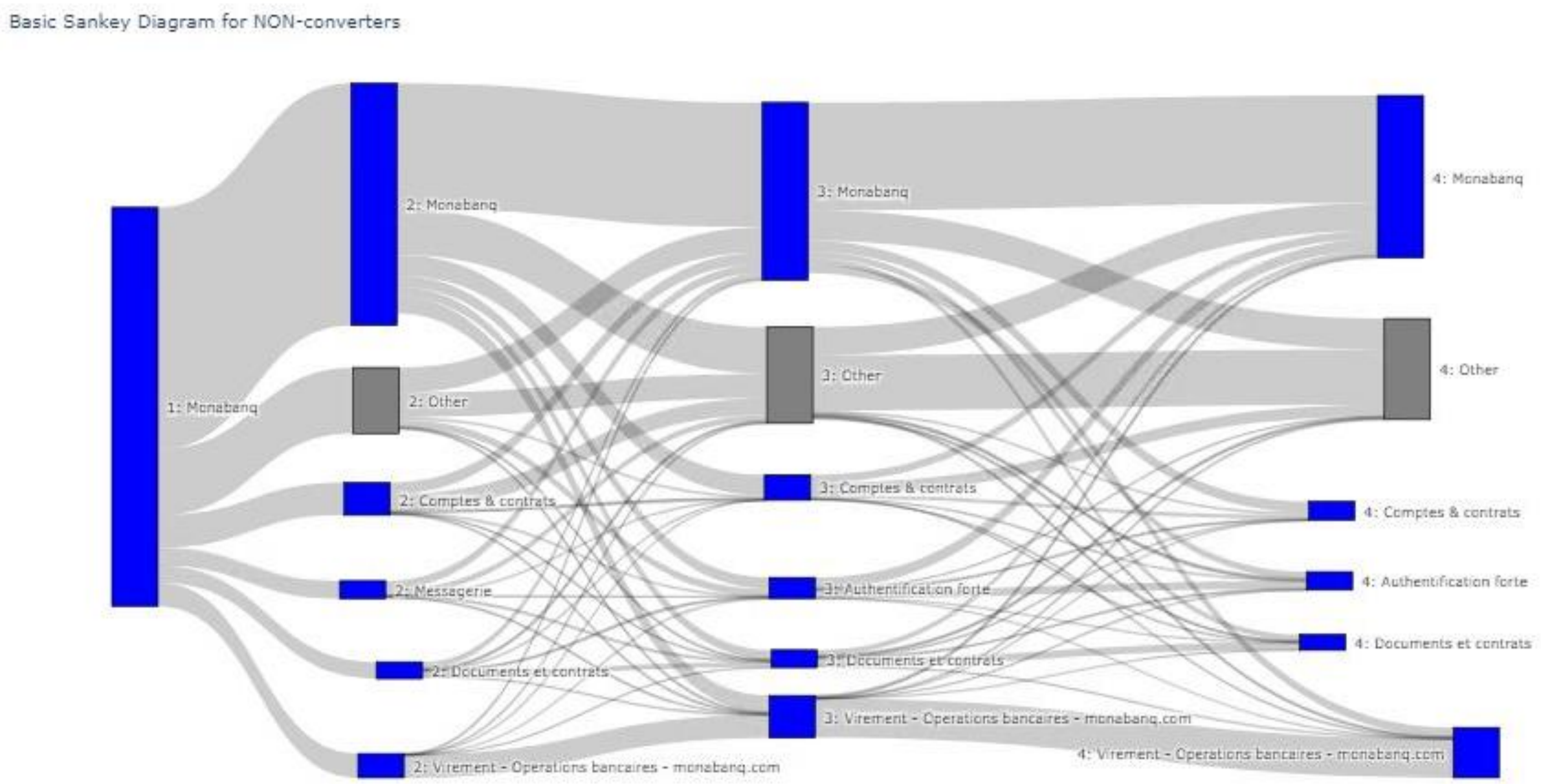
FOCUSED OUTCOME

BUYER PERSONA

IDEAL PATH

IÉSEG SCHOOL OF MANAGEMENT

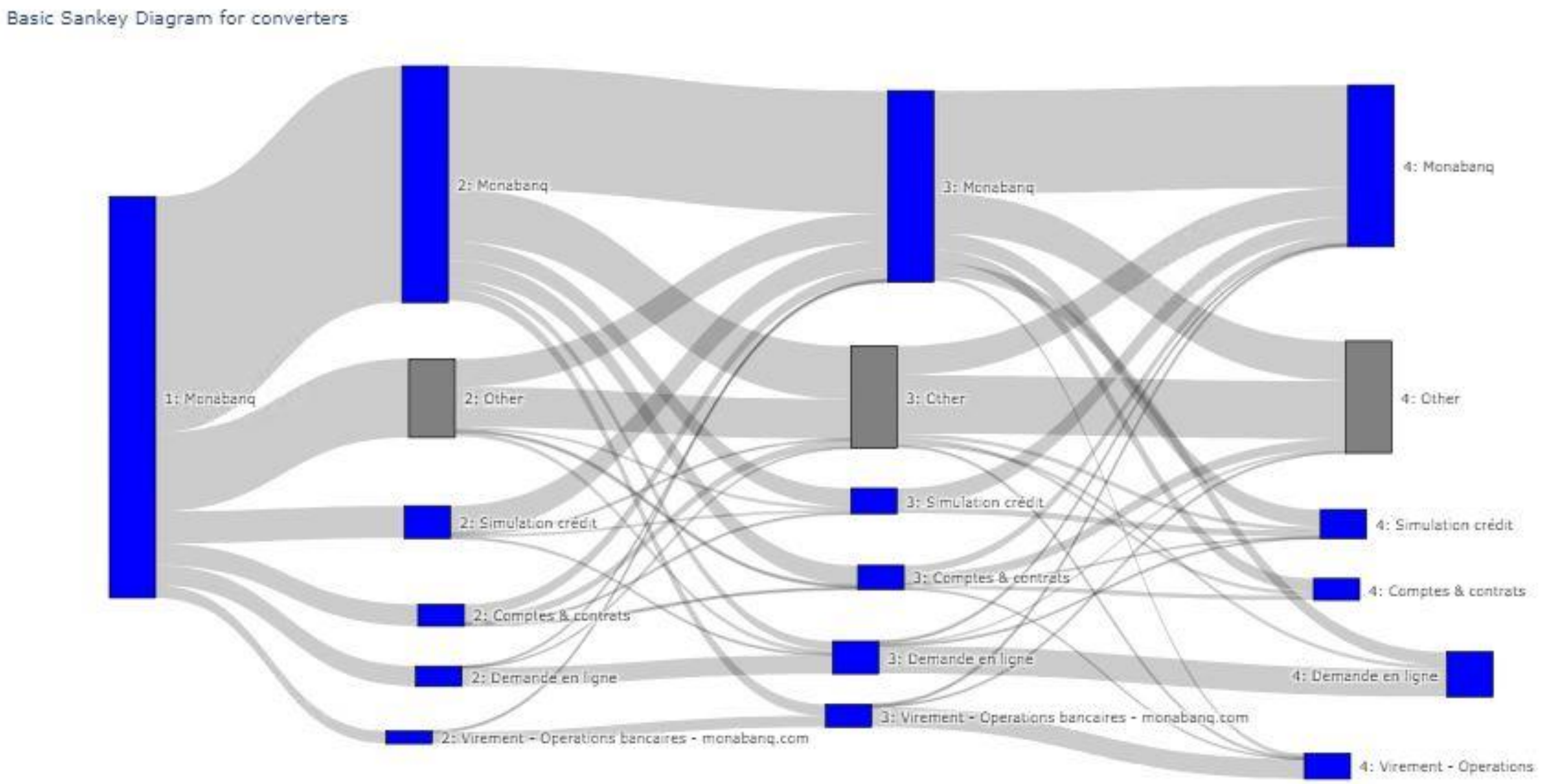
Navigation Paths followed by Non-Converters



Business Suggestion:

Why not performing A/B testing on these pages by putting more credit banners to boost interest?

Navigation Paths followed by Converters



Business Suggestion:

Go more in depth of pages visited and put banners in those pages. This will help take advantage of this niche to promote credit products.

Buyer Persona

Most of the Customers which went ahead and purchased fall the **age group of 26-49** years. Majority are mainly from **Particulier** customer segment users falling under the **medium Risk category** (C-). Further, **PC** was seen to be the most preferred Device and **even chose Insurance Product** along with Credit purchase.

IÉSEG SCHOOL OF MANAGEMENT

Ideal Journey Path

The Most Popular/Ideal website journey path that lead to Credit Purchase by the Customer was when the Customer previously visited “Simulation Credit”, “Comptes & Contrats” and “Demande en Ligne” previously.

Important Drivers to Credit Purchase

After analysing various Features and their correlation to the Target variable we found that the following were the top ones result to High Likelihood of the Credit Purchase by the Customer (below case is for week T1) –

IÉSEG SCHOOL OF MANAGEMENT

Key Business Takeaways

Journey Insights for the Customers who finally made a Credit Purchase Convert vs that Didn't and Dropped out –

- Converters mostly visited credit related pages from the beginning to the end of the sessions, along with the main account page
- The main navigation path is centred on:
  - Credit simulation
  - Credit subscription
  - Bank transfers
  - Contracts and accounts

- Non-converters mostly stay on the main Monabanq account page from the beginning to the end of the session
- The main navigation path is centred on:
  - Message inbox
  - Documents online
  - Bank transfers
  - Consulting contracts and accounts

IÉSEG SCHOOL OF MANAGEMENT

Key Business Takeaways

Behavioural data of the analysed so far throughout can be used by Monabanq Marketing Team to –

- *Optimise the Website*
- *Improve Customer Journey Experiences*
- *Better Interaction with the Visitors*

Ultimately, pushing Customers from "Awareness" to "Purchase" and loosing the least at the "Consideration" stage


Thank You