Executive Summary

- The dataframe used contains 395.937 instances and 44 features. 1.8% of the population studied have contracted a personal loan (the dependent variable).

- When the house is between 10-35 years old, with >8-year-old vehicles and clients who check bank details through their Bank App, the probability of contracting a personal loan increases.

- Comparing to the total number of clients in the population, the results obtained show that there are more clients from the Cataluña and Malaga areas.However, the success rate is the highest in the Northwest and Cataluña áreas (as shown in the previous graphic).

- 5 models were used to decide the one which better adjusted to the binary classification analysis (Logistic Regression, Decision Tree Classifier, Random Forest Classifier, XGBoost, Stack ensemble model).

- For the Business Plan, basic assumptions are created to be able to build a Business case(100k clients, average loan 6k€,margen per loan 4%, success predicted customers 22%).

- A potential outcome of 21% ROI (~350k €) per year could result if these customers are addressed: clients with <40% success rate will be contacted by WhatsApp. Between 40% and 80% by mail + phone. More 80% by mail + Branch 9.4k clients.

- 9.4k will be managed by Sales Representative. Explaining to them why this model has selected them and help them find the best manner to present it to the client is crucial. For the Local Features selection, the model used was LIME(Local Interpretable Model-agnostic Explanations).

- By using LIME, Sales Representatives can use this information(10-years-old car or if his house is 23) to offer clients the possibility to change their car without having to pay all upfront or to ask if hi/she is thinking on reforming any part of his/her house.

Although the initial information included in the dataset is real, in order to present this work and to comply with current regulations, all data has been anonymized.