Product Propensity Model For Banking Customers

Public Project Summary

Mariam Bazzi

Supervisor: Prof. Gábor Békés

MSc in Business Analytics

June 2021

Nádor u. 9. | 1051 Budapest | Hungary

Objective

The project involves building a propensity model that will predict the level of interest of customers in acquiring the product offered by the bank. The model is to serve as a pre-lead filter for marketing campaigns. Since the objective is to predict the level of interest a prospective client has towards acquiring the product, we initially have to provide the model with cases where people have and have not acquired those products. This will help the model distinguish between both cases, customers who have acquired a product and others who have not.

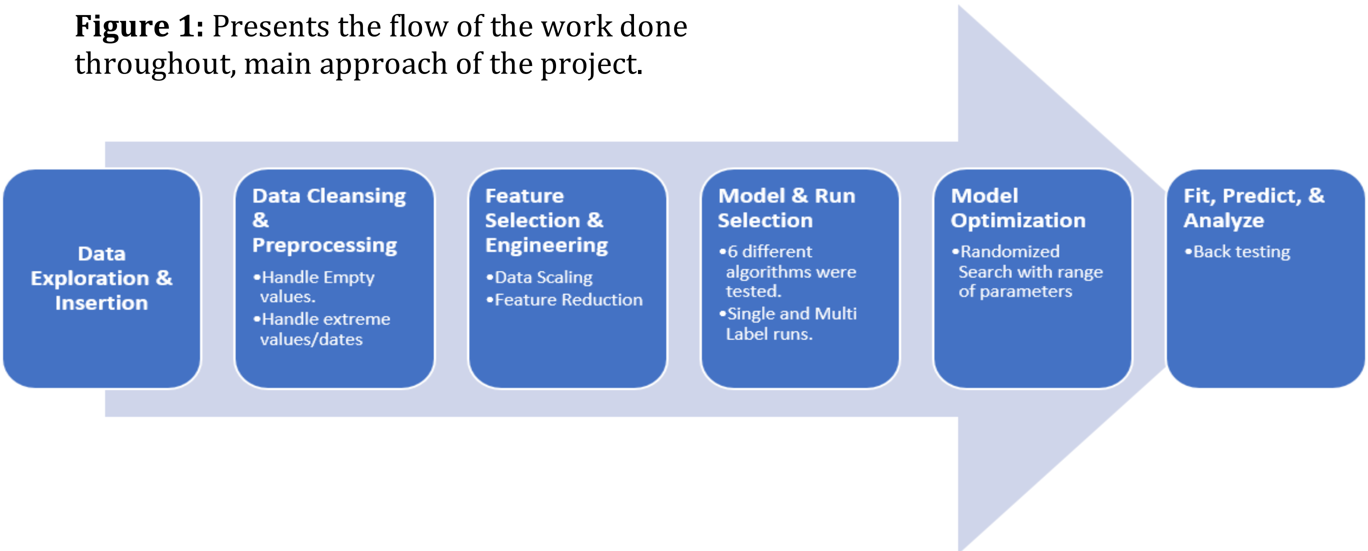
Background

Banks offer a suite of financial products and services to benefit potential customers. However, banks have, in the past, failed to adopt advanced techniques to determine the interest of potential customers who may require various financial products. Such scenarios ended up in losses due to low interest in financial products among these clients. In this regard, it is necessary for banking institutions to explore mechanisms to determine clients' interest in obtaining various financial products. The mechanism should analyze prospective customers' demographics and financials to assess their need in the products.

Methodology

The process can be summarized in six steps, the first being data exploration to analyze large data sets and uncover initial patterns and data points that are useful.

Figure 1: Presents the flow of the work done throughout, main approach of the project.



The second step was data cleansing and preprocessing to decide the most appropriate technique to handle extreme and null values as well as character encoding before inserting the data into the database. Feature selection and engineering was the third step, which primarily involved scaling data and reducing the features to ensure that features that are impactful for the model are only included. The fourth step was modeling, which entails picking previous occurrences with similar scenarios and applying the information to the model. This involved building single-class models per product, and a multi-class model for 3 products. This was proceeded with selecting the best performing model based on different performance metrics. Several algorithms were attempted to reach to the optimal tuned model. The algorithms attempted were Random Forest, XGBoost, AdaBoost Classifier, Naive Bayes, and Support Vector Classification. XGBoost was among the highest across the performance metrics that were computed. This was proceeded with model optimization, where the main activity was performing randomized searches for the optimal tuning parameters for each model. In this context, single-class per product runs proved most applicable as they could focus on each banking product separately. The multi-class model would require a significantly larger data set to provide more reliable results. The sixth step fitting the model, predicting and analyzing; whereby cross validation on the data extracted was run while fitting and back-testing the model on previous years served to validate the models.

Outcomes

The models were able to capture 70.7% of customers who actually acquired product 1 in 2020, while 53.4% of customers that acquired product 2 were successfully predicted. While analyzing the predictions extracted from the models, a critical point to compare is the conversion rate before and after the use of a model. Prior to the use of a model, the bank will traditionally be able to convert around 18.2% into acquiring product 1, and 39.7% into acquiring product 2. Stepping up, and moving towards the models' predictions, we can boost these rates to become 24.3% and 53.4% for product 1 and product 2 respectively. Both models have a similar growth in conversion rates when compared side by side at around 35% increase. This would provide a room to, initially, increase the revenue to the bank as the model will be able to capture and predict a significant number of clients that would be interested in acquiring a product when compared to previous means. And lastly, the bank will be able boost their profits as the they will require the same, if not less, resources and labor to deploy and execute a model which will provide more as well.

A benefit to the client is enabling banks to boost sales of their products and services. The key outcome would be realizing profits after identifying and targeting prospects that have a high interest in the financial products in both, customer and non-customer populations. Notably, banks started using these models to sell more by identifying clients who would need a specific product considering many variables such as demographics, balances and financials. The concerned bank would also avoid loss of financial resources on unfocused marketing activities that generate little or no benefits. A key outcome is determining the features that have higher impact and more significance to the outputs of the model. The outcome's significance is leading the bank to focus on prospects with these attributes.

Conclusion

The models consider many factors, including a client's demographics, products previously acquired, balances, and numerous financial attributes. The process was automated in order to improve modeling outputs such that they become holistic and reliable. These models were tailored for a direct use by the bank to identify the prospective clients that have a high interest in acquiring a product. Remarkably, the outcome culminated in a model that exhibits optimal performance regarding predicting clients' levels of interest in a financial product, and highlighted observations which are impactful on the performance of the model that the bank should concentrate on.