Banco Federal de Finanças

Marketing Campain Analysis

­­

## presented by

Lucas Nethercott

­William Figueroa

Austin Eldredge

Nicholas Balabanov

Elias Moreira

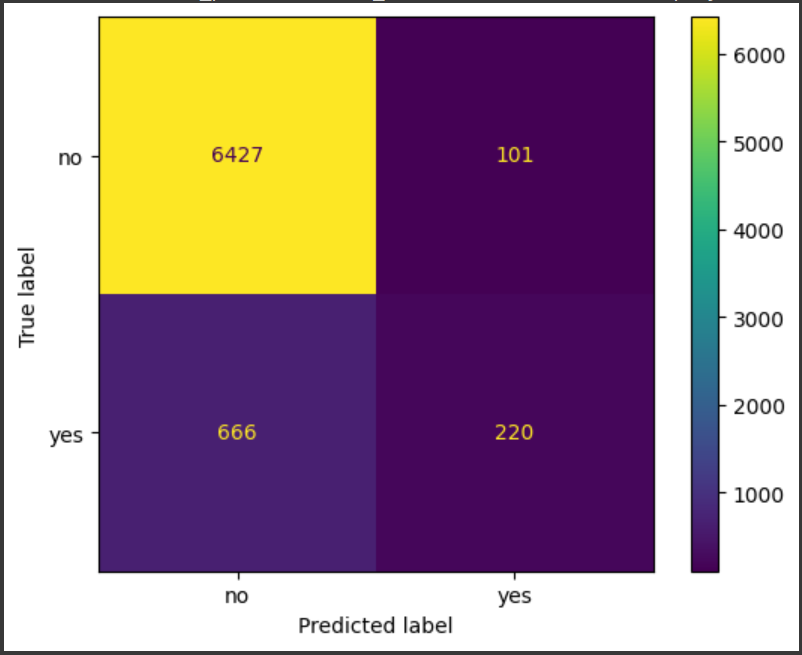
1. Background

This executive summary provides an overview of the background and objectives of the targeted telemarketing campaign undertaken by Banco Federal de Finanças in Aveiro, Portugal. The bank recently completed a major telemarketing campaign, which did not yield satisfactory results according to the management. As a result, the bank is looking to leverage machine learning and data science to improve the effectiveness of their marketing efforts.

The primary focus areas for the stakeholders include identifying interesting customer segments based on the previous marketing campaign, determining the types of customers most likely to respond positively to future campaigns, and proving the efficacy of a targeted campaign based on data science compared to randomly selecting customers.

To address these concerns, the data science team has been assigned specific tasks:

1. Create a machine learning model to identify customers who are most likely to subscribe to a term deposit
2. Find actionable patterns in the campaign results
3. Create additional models specific to social and economic indicators
4. Persist each model to it’s own file
5. Methodology
6. To address the task of identifying customers who are most likely to subscribe to a term deposit, we employed a Decision Tree Classifier model. The following steps outline our methodology:
7. Importing Libraries and Reading Data: We imported the necessary libraries, including pandas for data manipulation, train\_test\_split for splitting the data into training and test sets, and scikit-learn for the machine learning components. The bank dataset was read from a CSV file.
8. Preparing and Preprocessing the Data: The data was split into training and test sets, 80% for training and 20% for testing. The training and test data were prepared by separating the target variable ('y') from the features, which represented whether the given customer had subscribed to a term deposit. The features were then standardized.
9. Creating and Training the Model: We trained a Decision Tree Classifier model with a maximum depth of 5, using the preprocessed data. To streamline the workflow, we created a pipeline that combined the preprocessing steps with the decision tree model.
10. Evaluating the Model's Performance: The trained model was evaluated on both the training and test data. Predictions were made on the training data, and accuracy, classification report, and confusion matrix were computed to assess the model's performance. The same evaluation metrics were calculated on the test data to gauge how well the model generalized to unseen instances.
11. Testing the Holdout Data and Exporting Results: The holdout dataset, provided as a separate CSV file, was loaded, and predictions were made using the trained model. The predictions were saved in a DataFrame and exported as a CSV file. The file was then downloaded for further analysis and decision-making.

The Decision Tree Classifier model achieved a training accuracy of approximately 90.4% and a test accuracy of around 89.7%. The classification reports and confusion matrices provide additional insights into the model's precision, recall, and F1-score for both the training and test sets.

1. To identify actionable patterns in the campaign results, we followed the following methodology: **(waiting to hear back from Elias)**
2. We then created two additional models specific to consumer confidence index using almost the same strategy outlined in the first task, with one caveat. The first is a predictor trained on the features that are related to the bottom 50% of consumer confidence index rating and the second is a predictor trained on the features that are related to the top 50% of consumer confidence index rating.

A yellow and purple squares with numbers

Description automatically generated with low confidenceHigh consumer confidence index:

A picture containing text, screenshot, colorfulness, software

Description automatically generatedLow consumer confidence index:

1. Python Notebooks

Below is a link to our github repo with the notebooks we used during this case study:

<https://github.com/wmezadev/CSE-450-TEAM-4>