

Group Name: Bank Marketing Campaign

Members:

Sebastian Bucheli, smbucheli@gmail.com, Ecuador, Fundación Ecociencia, Data Science.

Nolan Piloza-Hibbit, npilozo@hotmail.com, Canada, University of Waterloo, Data Science

Problem description

There is information about a marketing campaign of a Portuguese bank and they have data of phone calls. The main objective is to predict -with the collected data- if a person will subscribe to the product.

Business understanding

Data bank almost always needs to find insights in imbalanced data. It is because the probability of one event is very low compared with the other events. Even so the banks try to identify patterns in those negative events and try several possible solutions.

In this project we are going to deal with imbalanced data. There are distinct approaches about dealing with imbalanced data and we are going to see two of them and analyze which one is better and why. Also, we have a big proportion of categorical variables compared with the numerical variables and we will need to make some transformations after the analysis (EDA) to perform the best model.

Project lifecycle along with deadline

Refer to Bank Marketing Campaign Lifecycle spreadsheet

Data Intake report

Refer to Bank Marketing Data Intake Report PDF

Github Repo link

<https://github.com/Sebasdel95/dataglacier-internship/tree/main/week7/deliverables>