

Classification Problem

Now its time to work on a Project based on a real life Business Problem.

Problem Statement

Your client is a retail banking institution. Term deposits are a major source of income for a bank.

A term deposit is a cash investment held at a financial institution. Your money is invested for an agreed rate of interest over a fixed amount of time, or term.

The bank has various outreach plans to sell term deposits to their customers such as email marketing, advertisements, telephonic marketing and digital marketing.

Telephonic marketing campaigns still remain one of the most effective way to reach out to people. However, they require huge investment as large call centers are hired to actually execute these campaigns. Hence, it is crucial to identify the customers most likely to convert beforehand so that they can be specifically targeted via call.

You are provided with the client data such as : age of the client, their job type, their marital status, etc. Along with the client data, you are also provided with the information of the call such as the duration of the call, day

and month of the call, etc. Given this information, your task is to predict if the client will subscribe to term deposit.

Data

You are provided with following files:

1. train.csv : Use this dataset to train the model. This file contains all the client and call details as well as the target variable “subscribed”. You have to train your model using this file.

2. test.csv : Use the trained model to predict whether a new set of clients will subscribe the term deposit.

Data Dictionary

Here is the description of all the variables :

Variable	Definition
ID	Unique client ID
age	Age of the client
job	Type of job
marital	Marital status of the client
education	Education level

default	Credit in default.
housing	Housing loan
loan	Personal loan
contact	Type of communication
month	Contact month
day_of_week	Day of week of contact
duration	Contact duration
campaign	number of contacts performed during this campaign to the client
pdays	number of days that passed by after the client was last contacted
previous	number of contacts performed before this campaign
poutcome	outcome of the previous marketing campaign
Subscribed (target)	has the client subscribed a term deposit?

How good are your predictions?

Evaluation Metric

The Evaluation metric for this competition is accuracy.

Solution Checker

You can use solution_checker.xlsx to generate score (accuracy) of your predictions.

This is an excel sheet where you are provided with the test IDs and you have to submit your predictions in the “subscribed” column. Below are the steps to submit your predictions and generate score:

- a. Save the predictions on test.csv file in a new csv file.
- b. Open the generated csv file, copy the predictions and paste them in the subscribed column of solution_checker.xlsx file.
- c. Your score will be generated automatically and will be shown in **Your Accuracy Score** column

You can also check out the baseline Python Notebook provided to get started.