FactElytics

Mode: CASE STUDY
Description: DATA ANALYSING & MACHINE LEARNING MODEL BUILDING

PROBLEM STATEMENT

Marketing campaigns are characterized by focusing on the customer needs and their overall satisfaction. Nevertheless, there are different variables that determine whether a marketing campaign will be successful or not. Some important aspects of a marketing campaign are as follows:
Segment of the Population: To which segment of the population is the marketing campaign going to address and why? This aspect of the marketing campaign is extremely important since it will tell which part of the population should most likely receive the message of the marketing campaign.
Distribution channel to reach the customer's place: Implementing the most effective strategy in order to get the most out of this marketing campaign. What segment of the population should we address? Which instrument should we use to get our message out? (Ex: Telephones, Radio, TV, Social Media Etc.)
Promotional Strategy: This is the way the strategy is going to be implemented and how are potential clients going to be addressed. This should be the last part of the marketing campaign analysis since there has to be an in-depth analysis of previous campaigns (If possible) in order to learn from previous mistakes and to determine how to make the marketing campaign much more effective.
You are leading the **marketing analytics** team for a banking institution. There has been a revenue decline for the bank and they would like to know what actions to take. After investigation, it was found that the root cause is that their clients are not depositing as frequently as before. Term deposits allow banks to hold onto a deposit for a specific amount of time, so banks can lend more and thus make more profits. In addition, banks also hold better chances to persuade term deposit clients into buying other products such as funds or insurance to further increase their revenues.
You are provided a dataset containing details of marketing campaigns done via phone with various details for customers such as demographics, last campaign details etc. Can you help the bank to predict accurately whether the customer will subscribe to the focus product for the campaign - Term Deposit after the campaign?

Data Description

Train Set

Train set contains the data to be used for model building. It has the true labels for whether the customer subscribed for term deposit (1) or not (0)

Variable	**Description**
id	Unique identifier for each sample in the dataset. Cannot be used for modelling
customer_age	Age of the Customer in years
job_type	Type of job of the customer
marital	Marital Status of the Custmer
education	Education Level of the Customer
default	Whether customer has Defaulted in Past
balance	Current Balance in the Customer's Bank
housing_loan	Has customer taken a Housing Loan
personal_loan	Has customer taken a Personal Loan
communication_type	Type of communication made by the bank with the customer
day_of_month	Day of month of the last contact made with customer
month	Month for the last contact made with customer
last_contact_duration	Last Contact duration made with the customer (in seconds)
num_contacts_in_campaign	Number of contacts made with the customer during the current campaign.
days_since_prev_campaign_contact	Number of days passed since customer was contacted in previous campaign.
num_contacts_prev_campaign	Number of contacts made with the customer during the previous campaign.
prev_campaign_outcome	Success or Failure in previous Campaign.
term_deposit_subscribed	(Target) Has the customer taken a term deposit ?

Test Set

Set of calls for which the prediction needs to be done regarding the subscription status of the customer for term deposit post campaign

Variable	Description
id	Unique identifier for each sample in the dataset. Cannot be used for modelling.
customer_age	Age of the Customer in years
job_type	Type of job of the customer
marital	Marital Status of the Customer
education	Education Level of the Customer
default	Whether customer has Defaulted in Past
balance	Current Balance in the Customer's Bank
housing_loan	Has customer taken a Housing Loan
personal_loan	Has customer taken a Personal Loan
communication_type	Type of communication made by the bank with the customer
day_of_month	Day of month of the last contact made with customer
month	Month for the last contact made with customer
last_contact_duration	Last Contact duration made with the customer (in seconds)
num_contacts_in_campaign	Number of contacts made with the customer during the current campaign.
days_since_prev_campaign_contact	Number of days passed since a customer was contacted in the previous campaign.
num_contacts_prev_campaign	Number of contacts made with the customer during the previous campaign.
prev_campaign_outcome	Success or Failure in the previous Campaign.

Sample Submission:

Format for making the submission for predictions on the test set

id: Unique id for each call

term_deposit_subscribed: whether term deposit was subscribed post call. (1/0)

Evaluation Metric

The evaluation metric for this competition is **binary F1 Score.**

Guidelines for Final Submission

Please ensure that your final submission includes the following:

1. **Excel file** containing the **predictions** of whether term deposit was subscribed.
2. **Code file** (convert as html file) for reproducing the submission, note that it is mandatory to submit your code for a valid final submission. Use R, python or any other language to build machine learning model.
3. **Your PowerPoint presentation file** where you will showcase your insights about data, machine learning model (include all the assumptions you took) and your business recommendations (if necessary).

NOTE:

· Make a zip folder for all three submission files and **upload ONLY zip the file**.
· Nomenclature for the zip file is **Drishti 2021_Team Name_College Name**.
· For Powerpoint presentation, ensure to wrap up all slides in maximum 10 slides (including Intro/ thankyou/appendix slides).

For any queries:
Anuvrat Shukla
Anuvrat.shukla@siom.in
8076317980