



BANK MARKETING CAMPAIGN

NAME: CATHERINE SANDA

EMAIL:sandacate@gmail.com

COUNTRY: KENYA

COLLEGE: JOMO KENYATTA UNIVERSITY OF AGRICULTURE AND TECHNOLOGY

SPECIALIZATION: DATA SCIENCE

GITHUB REPO LINK: <https://github.com/cate6495/Bank-week10>

PROBLEM DESCRIPTION

We are given data related to direct marketing campaigns i.e. phone calls of a bank in Portugal.

The classification goal is to predict whether a client will subscribe or not(yes/no) to a term deposit (variable y).

EXPLORATORY DATA ANALYSIS

This dataset had 21 columns and 41188 rows, there were no missing values.

TARGET VARIABLE

The dataset is highly imbalanced with majority of the data points belonging to the no class.

CATEGORICAL VARIABLES

JOB

Most of the people contacted by the bank had admin jobs .Very few people had unknown jobs.

Comparatively people with admin jobs subscribed more for the deposits compared to those with other jobs.

MARITAL

More than half of the people contacted were married and they subscribed more than people with any other marital status.

EDUCATION

The bank contacted people with university degrees the most.They did not contact illiterate people.

Equally a high percentage of those who had university degrees subscribed.Comparatively in terms of those who did not subscribe considering education with quite a number of people university degrees did not.

DEFAULT

People who had not defaulted in their credit were contacted the most by the bank. The bank did not contact people who had defaulted.

People with no as their default status are the most who have and have not subscribed.

HOUSING

People with housing loans have been contacted more by the bank followed by people with no housing loans, very few people with unknown status of housing loans were contacted.

Those with housing loans subscribed more.

PERSONAL LOAN

The bank mostly contacted people with no personal loans for the deposits.

Majority of the people with no personal loans did not subscribe. Very few people whose personal status loan is unknown subscribed.

MONTH

The bank contacted people mostly in the month of May, July, August and June, they did not contact people in the month of January and February.

Equally those people subscribed more in the month of May.

DAY OF WEEK

The bank contacts people on Monday, Tuesday, Wednesday, Thursday and Friday. Few people were contacted on Friday.

There number of people subscribing during each day are almost equal, same to those who do not subscribe.

CONTACT

Majority of the people contacted used cellular contact.

Cellular contact also led to higher conversions than telephonic point of contact.

POUTCOME

Majority of the people the bank contacted during this campaign had not been contacted in the previous campaign.

People with previous outcome as success have subscribed more for the deposits more than people who have not subscribed. People with non existent previous outcome subscribed more than any other.

NUMERICAL VARIABLES.

AGE

Data concerning age was mostly skewed to the right. There number of people who subscribed was more than those who did not subscribe at age 60+

DURATION

Most of the call duration for people who subscribed is between 0-2000 whereas duration with people who had not subscribed is between 0-1000

CAMPAIGN

There is so much overlapping in the campaign feature it might not be very useful in prediction.

PDAYS

Most of the pdays values is 999 which means most of the clients have not been contacted by the bank.

PREVIOUS

Most clients had not been contacted before and as such previous with value 0 doesn't help in predicting whether a client will subscribe for a term deposit or not.

EMP.VAR.RATE

Emp.var.rate is very helpful in prediction.

CORRELATION

Emp.var.rate, euribor3m, nr.employed and cons.price.index have very high correlation.

FINAL RECOMMENDATION

The most influential variables are duration,nr.employed,euribor3m and emp.var.rate

Duration has a positive effect of people saying yes since the longer conversation the more inclined the client will be to the term deposit.

Euribor3m denotes the Euribor 3 month rate which is an indicator of the average interbank interest rates.The higher the interest rates the more a client will be inclined to the term deposit.

Emp.var.rate which is the change of the employment rate has a negative impact of the term deposits since a change in the employment rate will make the clients less likely to subscribe to term deposits.

The bank should target the relatively old clients.

The bank should also prioritize those who were part of the previous campaigns.