

Business Understanding

LISUM33 – BankBytes

Model Targets:

The main objective at this stage is to identify key business variables that the analysis requires in order to perform predictions.

ABC bank wants to find whether there is any relationship between customer interaction and closing a sale. Since the previous marketing campaign were based on phone-calls and the outcome of a sale oftentimes required more than one phone-call to the same client, optimizing this process would mean finding the most responsive clients to phone-call marketing.

Therefore, the variables that represent the details of these phone-calls carry the most weight in understanding how to shortlist those customers for the next tele-marketing, SMS/email marketing campaigns whose chances of buying the product is more.

Project Goals:

The goal of the project is to find a demographic(s) of customers that are more likely to purchase ABC Bank's term deposit product.

Finding the demographics of customers is useful for the tele-marketing campaign as the ads can be tailored to target those groups of customers through visual ads, this can be achieved by clustering methods and could potentially target more than just one group in the same ad.

Regression methods can be used on the entire database to find all the individual customers (whose data the bank already holds) who are more likely to purchase the product, so that they can be targeted by the SMS/email marketing campaigns.

The tele-marketing campaign would aid in bringing in new customers via the new product or push existing customers to purchase the same product and the SMS/email marketing campaigns would aid in informing customers that there is a new service being offered.

Developing a model that accurately predicts whether the term deposit subscription was purchased would be the end goal of the project.

Data Sources:

The data source is the "Bank Marketing" dataset that can be found at this [LINK](#), this source contains the necessary data from ABC Bank about the call logs to customers and the conversion rates can be calculated based on regression modelling techniques. It also contains features that can be used for grouping demographics based on income, loans, education, jobs, etc.

The majority of this dataset does not have missing values, but factors like when the client was last contacted, if at all or how many times the client was contacted do not have missing values and provide us with a stable dataset to build a model(s).

Data Dictionary:

Input variables:

bank client data:

1. age (numeric)
2. job: type of job
(categorical: "admin.", "unknown", "unemployed", "management", "housemaid", "entrepreneur", "student", "blue-collar", "self-employed", "retired", "technician", "services")
3. marital: marital status (categorical: "married", "divorced", "single"; note: "divorced" means divorced or widowed)
4. education (categorical: "unknown", "secondary", "primary", "tertiary")
5. default: has credit in default? (binary: "yes", "no")
6. balance: average yearly balance, in euros (numeric)
7. housing: has housing loan? (binary: "yes", "no")
8. loan: has personal loan? (binary: "yes", "no")

related with the last contact of the current campaign:

9. contact: contact communication type (categorical: "unknown", "telephone", "cellular")
10. day: last contact day of the month (numeric)
11. month: last contact month of year (categorical: "jan", "feb", "mar", ..., "nov", "dec")
12. duration: last contact duration, in seconds (numeric)

other attributes:

13. campaign: number of contacts performed during this campaign and for this client (numeric, includes last contact)
14. pdays: number of days that passed by after the client was last contacted from a previous campaign (numeric, -1 means client was not previously contacted)
15. previous: number of contacts performed before this campaign and for this client (numeric)
16. poutcome: outcome of the previous marketing campaign (categorical: "unknown", "other", "failure", "success")

Output variable (desired target):

17. y - has the client subscribed a term deposit? (binary: "yes", "no")