**Bank Marketing Dataset**

**Data Dictionary**

|  |  |  |
| --- | --- | --- |
| Variable | Definition | Key |
| age | age of the customer. | numeric |
| age group | what age group does the customer lie | categorical |
| eligible | if the customer is eligible for the talk or not. | binary (yes, no) |
| job | type of job | categorical |
| salary | salary of the customer | numeric |
| marital | married or not? | categorical |
| education | level of education completed | categorical |
| marital-education | married or not- education. | categorical |
| targeted | if the customer or being targeted or not | yes or no |
| default | has credit in default? | categorical |
| balance | remaining balance in their accounts | numeric |
| housing | has housing loan? | categorical |
| loan | has prior personal loan? | categorical |
| contact | contact communication type | categorical |
| day | day of the month | numeric |
| month | last contact month of year | categorical |
| duration | last contact duration. highly affects output | numeric |
| campaign | number of contacts performed during this campaign and for this client | numeric |
| pdays | number of days that passed by after the client was last contacted from a previous campaign (999 means client was not previously contacted ) | numeric |
| previous | number of contacts performed before this campaign and for this client | numeric |
| poutcome | outcome of the previous marketing campaign | categorical |
| y | has the client subscribed a term deposit? | binary (yes, no) |
| response | response of the person (associated to *“y”* variable) | binary (1,0) |

**Objective**

Try to build a classification model that predicts if the customer will buy the product or not.

**Targets**

*“default”*

*“y”*

**Main Questions**

- **SMOTE** for “*default*” variable target

Resources:

<https://www.analyticsvidhya.com/blog/2020/10/overcoming-class-imbalance-using-smote-techniques/>

<https://machinelearningmastery.com/smote-oversampling-for-imbalanced-classification/>

**Preliminary Work**

- Cleaning and ordering of data

- Exploratory analysis

- Univariate Analysis

- Bivariate Analysis

- Multivariate Analysis

- Visualization

**Machine Learning**

- Classification Model