ISM6136.901S20 Data Mining

Data Mining Analysis of Bank Marketing Campaigns

Group A2
Team Members
Iretioluwa Bamgbala
Anu Yadav Kambalapally
Qianxi Liu
Hwei Hua Tan

Table of Contents

1. Introduction

Portugal Bank is the central bank of Portugese Republic. The increase of marketing campaigns decreased its impact on customers. Increasing competition, economic pressures, new technologies have caused companies to adapt to direct marketing techniques. So the companies are using Data Mining techniques to find the target group.

Business Problem :
The Portugal Bank has seen a great decrease in revenue of the bank and they want to know the cause of it. After an extended research they came to know that clients are not depositing as frequently as they used to do. One of the major ways the bank makes profits is through Term Deposit. Term deposit allows the bank to hold on to a deposit for a specified period of time so that they can invest in higher profit products to make profits. In addition to that, the bank can also retain these customers for a long time and encourage them to buy other products of the bank. Portugal Bank therefore wants to identify existing clients that have the maximum chances of subscribing to the term deposit so that they can target that particular group of clients.

As part of marketing techniques, Portugal Bank conducted campaigns based mostly on direct phone calls, offering bank clients to place a term deposit. If after all marketing efforts, clients had agreed to place a term deposit, the effort will be deemed successful. The main marketing effort has been made based on phone calls. Often, more than one contact to the same client would be required.

2. Data Set Exploration

The data set was found on Kaggle and was based on "Bank Marketing" UCI data set. There are 41,188 instances and 21 attributes in the data set. The 21 attributes included our output variable (y). Below is the description of all the attributes:

Output variable (desired target):

Target Variable is specified as Yes if the customer agrees to subscribe to term deposit and No if the customer disagrees.The Yes and No percentages are shown as below in the graph.

Subscription Status
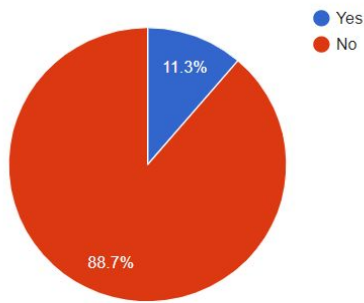
Variable	Description	Type of Variable
y	Has the client subscribed to a term deposit?	Binary: YES, NO

Input variables (bank client data):

Variable	Description	Type of Variable
age	Age of client	Numeric
job	Type of job	Categorical : admin, blue-collar, entrepreneur, housemaid, management, retired, self-employed, services, student, technician, unemployed, unknown)
marital	Marital status	Categorical: divorced, married, single, unknown; note: "divorced" means divorced or widowed
education	Client's education	Categorical: basic.4y, basic.6y, basic.9y, high.school, illiterate, professional.course, university.degree, unknown
default	Has credit in default?	Categorical: yes, no, unknown

| housing | Has a housing loan? | Categorical: yes, no, unknown |
| loan | Has a personal loan? | Categorical: yes, no, unknown |

Input variables (related with the last contact of the current campaign):

Variable	Description	Type of Variable
contact	Contact communication type	Categorical: cellular, telephone
month	Last contact month of year	Categorical: jan, feb, …., dec
dayofweek	Last contact day of the week	Categorical: mon, tue, wed, thu, fri
duration	Last contact duration, in seconds	Numeric

Input variables (other):

Variable	Description	Type of Variable
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	Numeric; 999 means client was not previously contacted
previous	Number of contacts performed before this campaign and for this client	Numeric
poutcome	Outcome of the previous marketing campaign	Categorical: failure, nonexistent, success

Input variables (social and economic context):

Variable	Description	Type of Variable
emp.var.rate	employment variation rate - quarterly indicator	Numeric
cons.price.idx	consumer price index - monthly indicator	Numeric
cons.conf.idx	consumer confidence index - monthly indicator	Numeric
euribor3m	interest rate at which a panel of banks lend money to one another with a maturity of 3 months	Numeric
nr.employed	number of employees - quarterly indicator	Numeric

3. Methodology - Decision Tree

We analysed the data set in SAS Miner using decision trees to find out the target group so that the bank can contact the specific group and make the customers subscribe to the term deposit.


Below is the image showing the results of the Decision tree and the number of times the decision tree used the variable to split.

Variable Name	Label	Number of Splitting Rules	Importance
duration		9	1.0000
nr_employed	nr.employed	1	0.8366
month		2	0.3344
pdays		2	0.2489
euribor3m		4	0.2215
poutcome		1	0.1449
cons_price_idx	cons.price.idx	1	0.0803
default		1	0.0742
contact		1	0.0685
previous		1	0.0677

As we can see, the variable **duration** was used 9 times to split and it has an importance of 100%
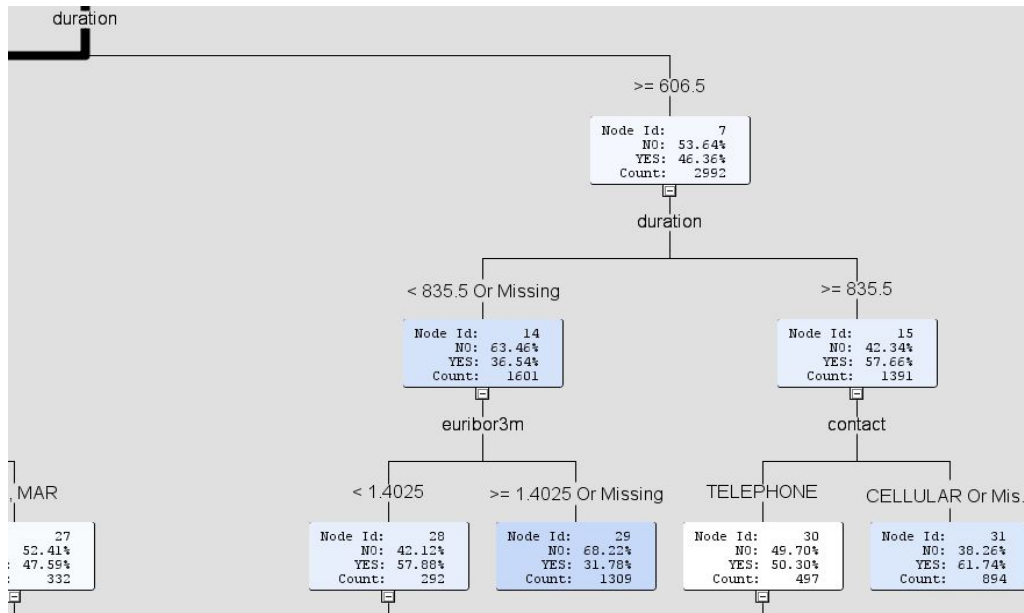
## 4. Results

Following is the result of decision tree:
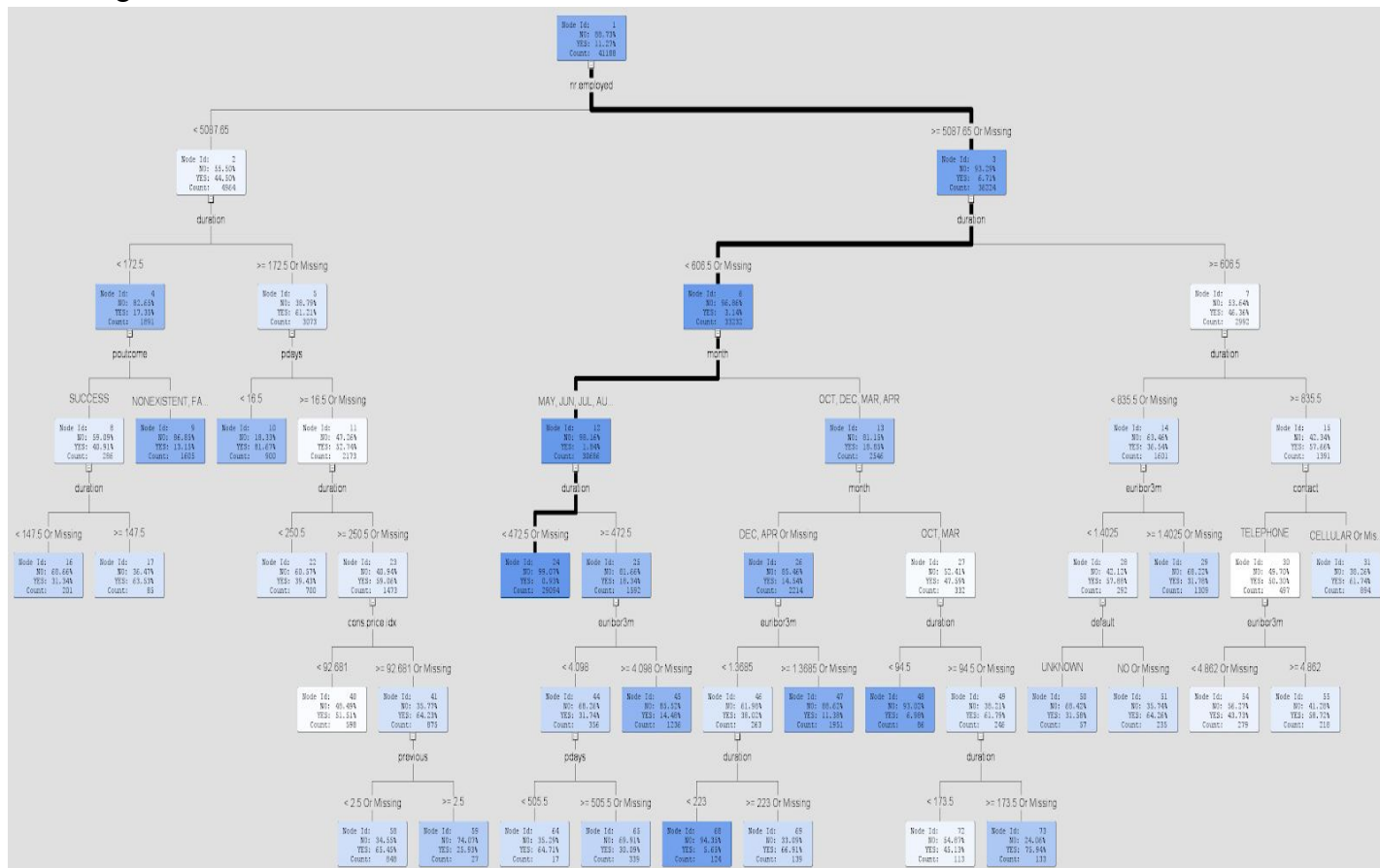
5. Evaluation

By evaluating the decision tree we found out the conditions upon each variable when the subscription rate is high.
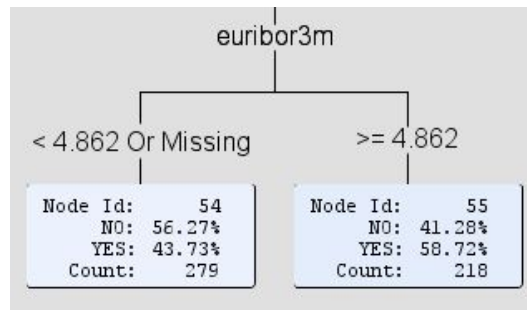
1. When the duration of the call made is greater than 835.5 seconds/14 minutes, the chances of the customer subscribing are high.

2. When customers are contacted in October, December, March and April, the chances of them subscribing are high.
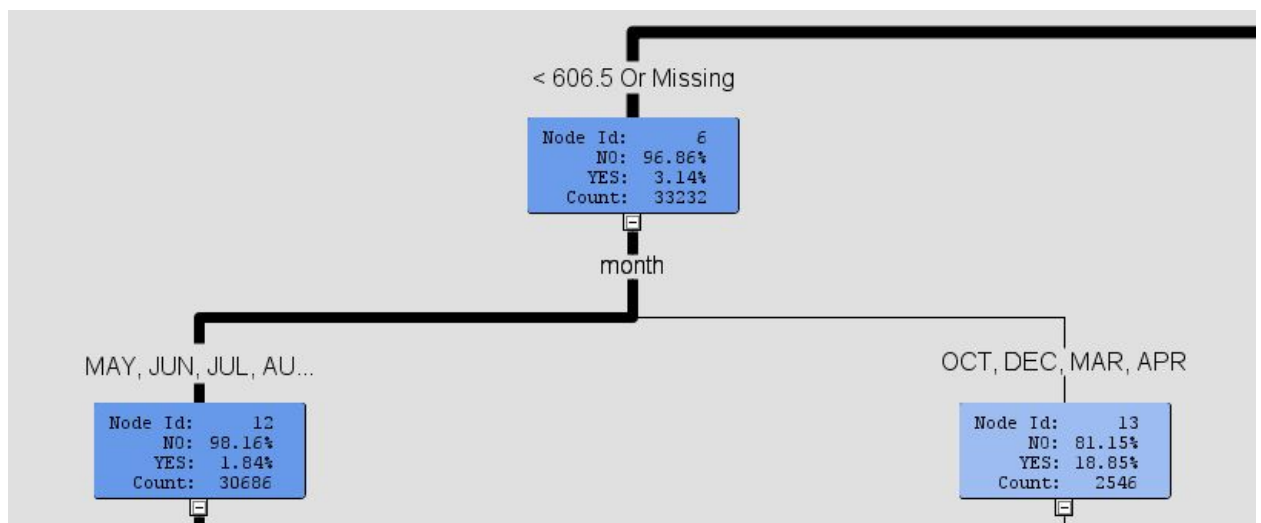
3. When the interest rate at which a panel of banks lend money to one another with a maturity of 3 months is greater than or equal to 4.862, the chances of subscription are higher.


```
                        euribor3m

        < 4.862 Or Missing          >= 4.862

     Node Id:      54            Node Id:      55
         NO:  56.27%                 NO:  41.28%
        YES:  43.73%                YES:  58.72%
      Count:      279             Count:      218
```

6. Recommendation

The business needs to optimise the Marketing Campaign
With the classification model and results from Decision Tree-- the conditions to predict the client subscribing a term deposit.

- Last contact duration should be greater than 835.5s/14mins.
- Mode of contact should be Cellular
- Contact month to be Oct,Dec,Mar,Apr
- Euribor rate 3 month should be greater than 4.862
- Pdays less than 16.5 days.