*ARPITA SHRIVAS*

*KAOUTAR ASSAH*

*SHRUTHI CHEGURI*

*ASHIRA CHUGH*

*MSIS 685 Architect, Build and Manage Big Data Applications*

*By Prof. Jean-Pierre Kuilboer*

**DESCRIPTIVE ANALYTICS**

Text

Description automatically generated

Graphical user interface, table

Description automatically generated with medium confidence

Graphical user interface, text, application, email

Description automatically generated

Table

Description automatically generated

Table

Description automatically generated

Table

Description automatically generated

Graphical user interface, text, application

Description automatically generated

Graphical user interface, text, application, email

Description automatically generated

**PRESCRIPTIVE ANALYTICS**

Table

Description automatically generated

Table

Description automatically generated

Table

Description automatically generated

Graphical user interface, text, application

Description automatically generated

Graphical user interface, text, application, email

Description automatically generated

The accuracy score of Random Forest Classifier looks better than logistic regression model.

**GRAPHS**

Histograms plots based on age

Chart, histogram

Description automatically generated

Chart, histogram

Description automatically generated

Chart, histogram

Description automatically generated

Chart, histogram

Description automatically generated

Application

Description automatically generated with medium confidence

Analysis of Day’s

Most of the calls made during 14-21st of the Month, and then some between 5-8th and some during 28-30th of the month.

Chart, bar chart, histogram

Description automatically generated

Analysis of Job’s

So, most of the customers are from Blue-collar, Management, technician and admin jobs.

Chart

Description automatically generated

Analysis of poutcome (Previous campaign outcome)

Most of the Previous campaign outcomes are unknown, so better data collection is needed.

Graphical user interface

Description automatically generated

**ANALYSING THE MARKETING RESULTS**

Graphical user interface, text, application

Description automatically generated

Conclusion: The data is related with direct marketing campaigns of a Portuguese banking institution. The marketing campaigns were based on phone calls. As mentioned, more than one contact to the same client was required, in order to access if the product (bank term deposit) would be (or not) subscribed so we are calculating the success of a Portuguese bank campaign based on the people. We extracted data to understand various data fields like marital status, gender, defaulter, etc. Further, we worked on regression models to check how effective are the given data fields on marketing campaign. We could see the model is effective as both logistic regression and random forest shows accuracy >85. We carried out visualizations and found that in current campaign 12% clients subscribed hence better targeting is required. Operational resources can be decreased by improving prediction in future of who is going to subscribe and who is not. Old customers who were previously contacted are more likely to subscribe than contacted for first time. We also tried to analyze the number of people last contacted by setting a threshold, which also shows that the customers are uniformly distributed by week4.