Data Science – Banking domain assessment

Domain:
Banking services data set

About:
DB Bank is a large public sector bank which has branches across the cities. It provides various services like savings account, current account, term deposits, personal loans, home loans etc. to customers. Whenever the bank conducts marketing on its new schemes, it will keep track of data related to customers' personal, social and economic details. Also, it maintains the detailing on efforts made to achieve success in the campaign.

Recently, the bank has conducted a campaign to market their term-deposit scheme. Campaigns was conducted based mostly on direct phone calls, soliciting the bank's customers to place a term deposit. After all the marking efforts, if the client had agreed to place a deposit, then the campaign is success, otherwise not (Target variable marked 'yes', or 'no').

Challenges:
It is a challenge for bank officials to target right people for a successful campaign. The marketing team must analyse various details like profession, income, age, education etc. of the customers. Also, the possibility of any existing loans, credit history etc. will give economic status of the customer and hence helps in understanding whether the person is capable of depositing money into term-deposit scheme of the bank.

The first step of this analysis is to assess what data is available and perform some exploratory and descriptive analytics to identify interesting and useful patterns, trends, and insights.

And the next step is to build a predictive model on the give data.

What is Expected?
Being a data analyst, you must come up a first step document that lists output of your exploratory analysis, any issues or problems you may see with data that need follow up, and some basic descriptive analysis that you think highlights important outcomes/findings from the data. Based on your findings, the next level of analysis will be charted out.

Also, you need to build appropriate predictive model for classifying successful and unsuccessful campaigns. You can perform comparative study of several predictive models with various approaches and give your inferences accordingly.

Here are some indicative types of analysis you can perform. Please note that this is not an exhaustive list, you may add more.

Data Preparation and Exploratory Data Analysis:

- Come up with appropriate results and visuals for the following:
    - Analysis of percentage turnout of marketing campaign
    - Right mode to contact the customers (telephone or mobile)
    - Analysis on attempts made to turn a person into successful depositor
    - Personal data analysis on marital status, existing loans, education, profession etc. and its impact on the campaign's success.
    - Socio-economical analysis of the customers
- Prepare the data by handling missing values, outlier analysis, data transformation and normalization.

Predictive Analysis:

- Build appropriate predictive model/s on the data.
- Compare various predictive models with appropriate regularization and/or hyper-parameter tuning.
- Evaluate the performance of the model.
- Identify the right metric to evaluate the performance of the model.
- Identify issues and concerns on the given data and suggest the best technique/s to overcome the issues.

Recommendations:

- As a data analyst, what are the approaches do you suggest the marketing team to identify ideal target group to make the campaign successful? Recommend based on your analysis.

NOTE: Results and graphs must be backed with appropriate inferences and insights.

Data Dictionary:

Bank client data:

- custAge: Age of the customer.
- profession: type of job
- marital: marital status
- schooling: Educational qualification
- default: has credit in default?
- housing: has housing loan?
- loan: has personal loan?

Data related with the last contact of the current campaign:

- contact: contact communication type
- month: last contact month of year
- day_of_week: last contact day of the week
- campaign: number of contacts performed during this campaign and for this client (includes last contact)
- pdays: number of days that passed by after the client was last contacted from a previous campaign (999 means client was not previously contacted)
- previous: number of contacts performed before this campaign and for this client
- poutcome: outcome of the previous marketing campaign

Data related to social and economic context attributes

- emp.var.rate: employment variation rate - quarterly indicator
- cons.price.idx: consumer price index - monthly indicator
- cons.conf.idx: consumer confidence index - monthly indicator
- euribor3m: euribor 3 month rate - daily indicator
- nr.employed: number of employees - quarterly indicator

Target variable:

- responded - has the client subscribed a term deposit?