Acquisition Analytics Case Study

Problem:
Predict the probability of a response from each prospect and target the ones most likely to respond to the next telemarketing campaign. The steps were as follows:

- Identify relevant predictor variables for a response using EDA.
- Build predictive models and choose the best one.
- Sort the prospects in order of decreasing probability of response (predicted by the best model) and target the top X% (or top Y deciles), where X would be determined by your business objective (e.g., maximising the overall response rate/number of responders at a fixed marketing cost).

Task
Set the business objective to achieving 80% of total responders at the minimum possible cost. The total number of responders is the total number of prospects who responded, from the available data of about 45,000 prospects.

Calculate the X in the top X%, i.e., how many prospects should be called to meet the business objective.

1. We found relevant variables using EDA which are:
   1.1 Age
   1.2 Job
   1.3 Marital
   1.4 Education
   1.5 Previous Default
   1.6 Housing
   1.7 Loan
   1.8 Contact
   1.9 Month
   1.10 Day of the week
   1.11 Campaign
   1.12 Gap after the last contact
   1.13 Previous
   1.14 Employment Variation Rate
   1.15 Consumer Price Index
   1.16 Euribor three-month rate
   1.17 No. of employees
2. We have performed modelling using Logistic Regression and used RFE for variable selection to select out of 15 variables and our result as follows:
   Impt variables predicted:
   Month: March, May, June, August, November
   Day of Week: Monday
   Job: Student
   Contact: Telephone
   Euribor three month rate
   Pdays: First time contacted

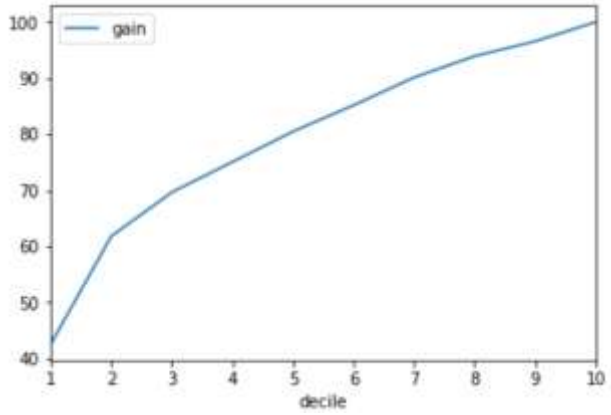
Our results:

Optimal cut off value: 0.1
Accuracy: 78%
Sensitivity(TP): 67%
Specificity(TN): 79%

3. Dataframe with the variables prospect ID, actual response, predicted response, predicted probability of response, duration of the call in seconds and cost of the call:

	Actual_response	Predicted_probablity	duration	Predicted_Response	Prospect_ID	duration_mins	cost
39340	1	9.067330e-01	101	1	139340	1.683333	2
39336	1	9.066381e-01	136	1	139336	2.266667	3
39258	1	9.060671e-01	215	1	139258	3.583333	4
40450	1	9.020609e-01	1064	1	140450	17.733333	18
40365	1	8.925504e-01	178	1	140365	2.966667	3
39153	1	8.917506e-01	363	1	139153	6.050000	7
39334	1	8.906684e-01	255	1	139334	4.250000	5
39255	1	8.900116e-01	290	1	139255	4.833333	5
40278	1	8.884697e-01	429	1	140278	7.150000	8

4. We can attain 80% of total conversions by targeting only top 50% of the total client base.
   Average call duration for targeting the top 80% prospects is 4.42 minutes
5. Lift chart:-
   x-axis should show the number of prospects contacted;
   y-axis should show the ratio of the response rate using the model and the response rate without using the model



6. Cost of acquisition:
   Cost incurred for acquiring 80% of customers using the predictive model = $ 101434