Acquisition analytics

- By Geetakrishnasai

Overview

- Objective
- Methodology
- Model metrics
- Results
- Conclusion

Objective

- Problem Description
  - Bank has run a telemarketing campaign previously and it has data about few attributes and corresponding response of the people
- Objective
  - Achieving 80% of total responders at the minimum possible cost

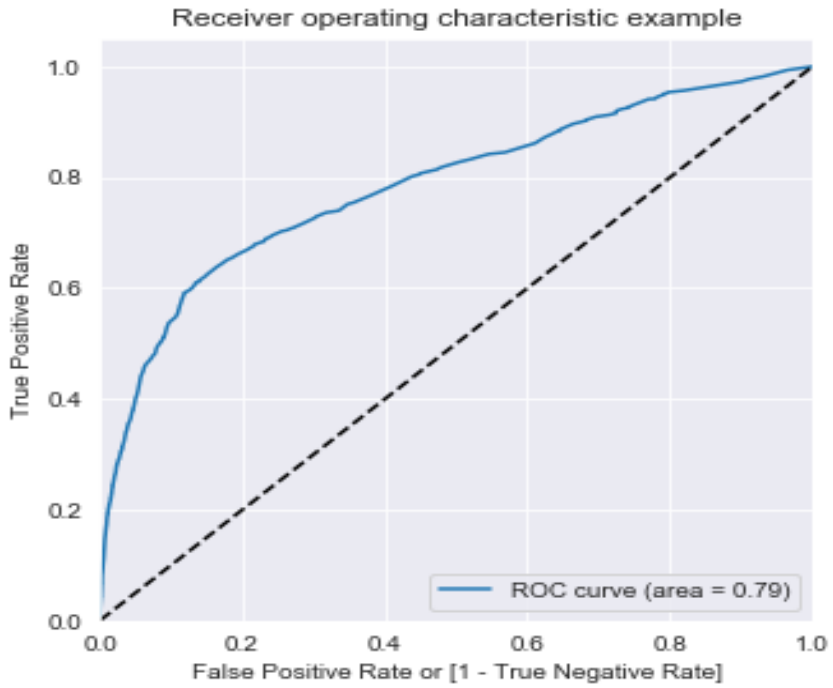
Methodology

- Performing EDA and identifying the important metrics
- Model building and hyper-parameter tuning
  - RFE, p-value and VIF values were used for deciding on important features
  - Logistic regression and grid search with AUC as performance criteria was used
- Choosing optimal cutoff value based on sensitivity and specificity curves
- Evaluating key metrics for decision making
  - Identifying number of deciles for reaching target of 80% of responders
  - Calculating average call duration for the identified deciles
  - Calculating the cost saving achieved using the model

Model Metrics : ROC

- **Final model equation**

- 0.42 + 0.392 x job_student -0.252 x contact_telephone +0.523 x month_dec + 0.834 x month_mar -0.956 x month_may-0.189 x month_nov -0.233 x day_of_week_mon + 0.424 x previous_Never contacted + 1.984 x poutcome_success + 0.182 x cons.price.idx-0.924 x euribor3m

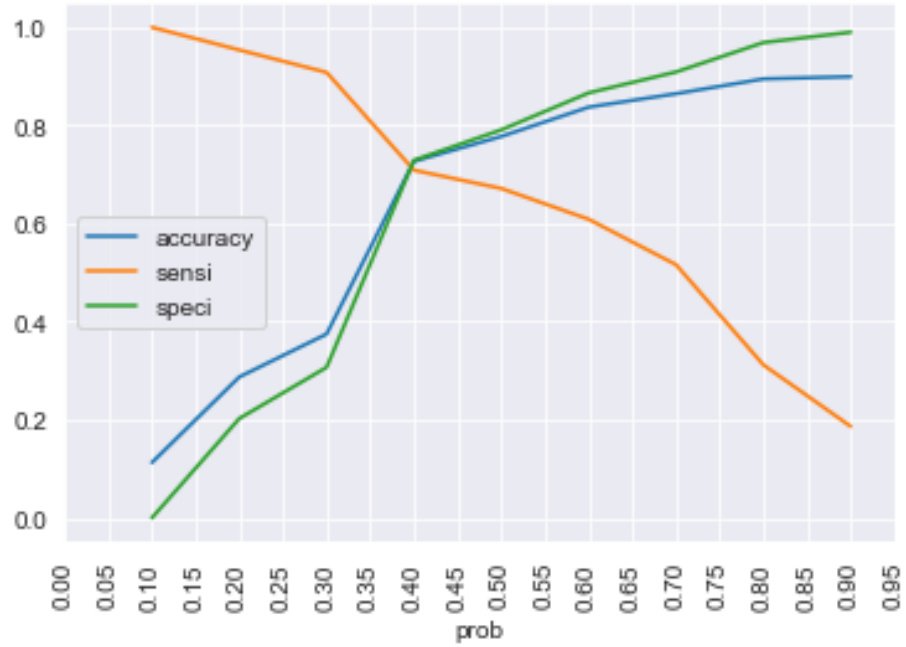


ROC curve with area of 0.79

Model Metrics : Optimal Cutoff

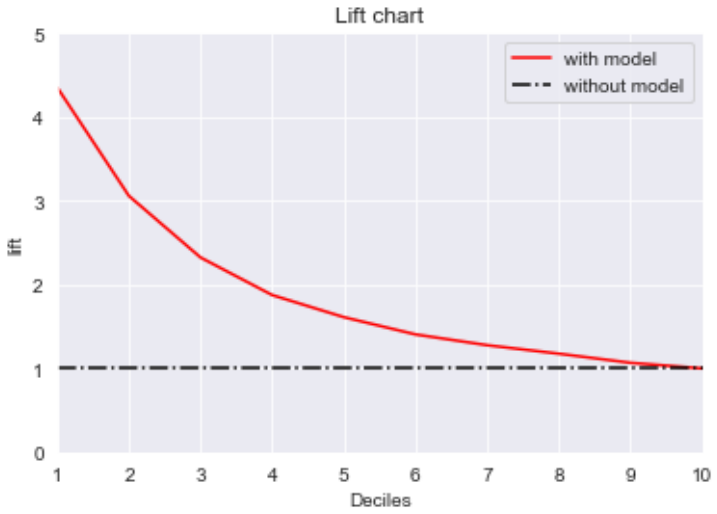
- The optimal cutoff probability is 0.39 , which is obtained using specificity and sensitivity curves (figure on the right)

- Accuracy – 0.72
- Sensitivity – 0.70
- Specificity – 0.73

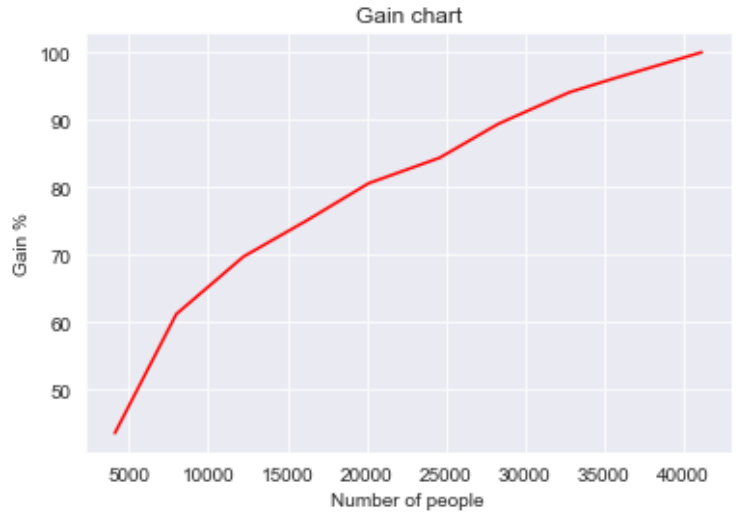


Results : Top X% Analysis

- Our objective was to target top 80% of total responder
- It can be achieved by targeting people in **top 5 deciles** (~ 80.6%)
- The **average call duration** for top 5 decile is **268 secs** (~ 4.5 mins)



Lift chart : Red line represents the model and dashed line represents no model



Gain chart : with the model

Results : Cost Analysis

- Cost analysis is performed based on both number of contacts and duration of call
- Assumptions:
  - Number of contacts made is considered as sum number of contacts made to each person
  - Average call duration is considered for computing call costs and it is assumed that it remains same across all the contacts made with a person
  - Monetary cost of call is considered to be 1 paisa per second

Based on Number of contacts made

Cost without model :

- Total number of contacts = 102971

Cost with model :

- Contacts required for top 5 deciles = 45318

Cost reduction achieved = 56%

Based on duration of call

Cost without model :

- Total number of contacts = 102971
- Avg call duration = 254 sec
- Cost (in Rs) = 2,61,905

Cost with model :

- Contacts required for top 5 deciles = 45318
- Avg call duration (top 5 deciles) = 272 sec
- Cost (in Rs) = 1,21,586

Cost reduction achieved = 54% (~ 1.4L in Rs)

Conclusion

- From the equation we can observe that
  - **poutcome_success, month_mar** have most **positive** effect on response variable
  - **month_may, euribor3m** have most **negative** effect on response variable
- The **average call duration** for targeting 80% of responders is **268 seconds** (~ 4.5 minutes)
- Acquisition Cost:
  - The **number of contacts** required to make reduced by **~57k** (about **56%**)
  - The **cost of calls** made reduced by **1.4L** (about **54%**)

- We have achieved a cost reduction of about 54% using the model