Lead Scoring Case Study

Group Members

Anagha
Anubhab

Problem statement

- Organization X, sells online courses from working professionals.
    - They gather customer details when they land on their website - leads
    - They use search engine and advertisements on other websites to gather leads
- They currently have a 30% a conversion rate (Lead - purchase of course)
    - If they receive a 100 leads a day, they convert only 30 of them
- Improve efficiency of the process by identifying high potential and low potential leads so that the relevant teams can allocate resources accordingly to increase conversion rate.
    - Identify "hot leads" by assigning a lead score to the lead based on customer data
    - Achieve a **80%** conversion rate at the end of the improvisation in process and sorting of leads

Methodology Followed

- Importing the Raw Data
- Cleaning of data and Data manipulation/preparation
  - Dropping unnecessary columns, columns with high number of N/A
  - Filling N/A with appropriate values
- Handling outliers:
  - By dropping the outliers or replacing with the median
- Segmentation
  - Separating the continuous variables and categorical variables
- Perform Univariate and Bivariate analysis
- Converting categorical variables into dummy variables
- Splitting into train and test set
- Rescaling variables

Methodology Followed

- Using RFE to select only the top features

- Running RFE with 20 variables as desired output

- Model building

  - Removing variables one-by-one, by checking their P-Value and VIF

  - P-Value less than 0.02 and VIF less than 5

- Predicted on the test data set:

  - Overall accuracy = **82%**

  - Sensitivity = **70%**

  - Specificity = **90%**

ROC Curve
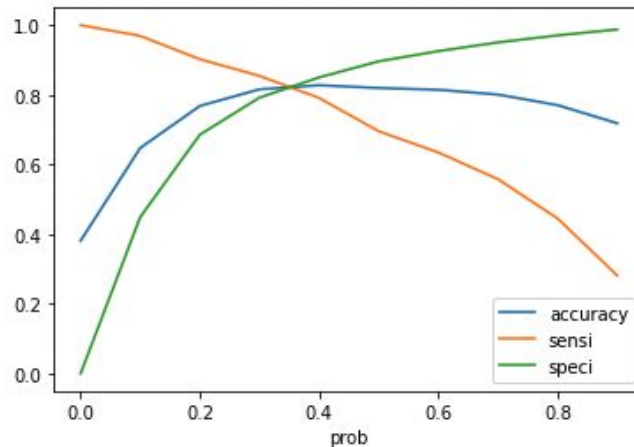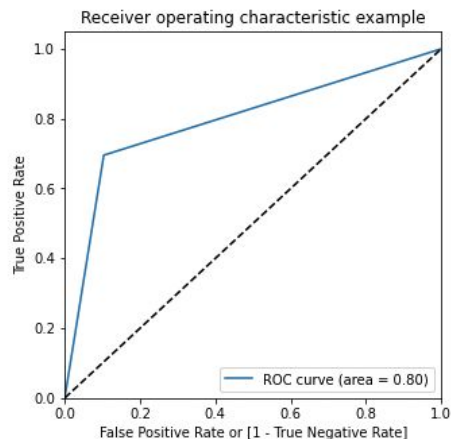
- *We can see that **0.32** is the optimum point to take it as a cutoff probability, since this is the point where we see Specificity & Sensitivity being balanced*

Conclusions

It can be concluded that the variables that mattered most in getting potential students are:

- Lead Profile_Unknown

- Lead Source_Olark Chat

- Last Activity:
    - SMS Sent
    - Olark Chat Conversation

- Total Time Spent on Website

THANK YOU