Summary

X Education wants out help to select the most promising leads, i.e. the leads that are most likely to convert into paying customers. The company requires us to build a model wherein you need to assign a lead score to each of the leads such that the customers with a higher lead score have a higher conversion chance and the customers with a lower lead score have a lower conversion chance. The CEO, in particular, has given a ballpark of the target lead conversion rate to be around 80%.

Steps Followed:

So to summarise, the steps that were performed throughout the model building and model evaluation were:

1. Data cleaning and preparation included the data inspections and removing the data with more nulls and treating the missing values

2. Handling categorical variables: this activity included Mapping categorical variables to integers and Dummy variable creation for these categorical variables

3. Then we did Test-train split and scaling of the data set

4. After that Model Building was started wherein, we did Feature selection using RFE (Coarse Tuning) and Manual feature elimination (using p-values and VIFs)

5. To test if the model is correct or not, we did Model Evaluation using below parameter

   - Accuracy

   - Sensitivity and Specificity

   - Optimal cut-off using ROC curve

   - Precision and Recall

6. And the final step was the Predictions on the test set

Below are the Final Observation done after executing above steps

Cut off was selected as 0.358 and we got the below which is giving us good accuracy and sensitivity and Specificity.

Evaluation Metrics for the train Dataset:-

- Accuracy :0.80
- Sensitivity:~0.80
- Specificity:0.81

- Precision: 0.72
- Recall: 0.80

Evaluation Metrics for the test Dataset:-

- Accuracy : 0.80
- Sensitivity: ~ 0.80
- Specificity: 0.81

Final Equation

(const   *-0.1437)+(Do Not Email   *-1.7971)+(Total Time Spent on Website*1.1116)+(Lead Origin_Landing Page Submission*   -1.0076   )+(Lead Origin_Lead Add Form*3.0175)+(Lead Source_Olark Chat*1.234)+(Lead Source_Welingak Website   *2.5089)+(Last Activity_Converted to Lead *-1.2517   )+(Last Activity_Had a Phone Conversation*2.6226)+(Last Activity_Olark Chat Conversation   *-1.3709   )+(Last Activity_SMS Sent *1.2727)+

(Last Activity_Unsubscribed*1.4019) + (What is your current occupation_Housewife*22.6448) + (What is your current occupation_Others*-1.2178)+(What is your current occupation_Working Professional*2.4015)+(Specialization_Others   *-0.9283)