Vu Phuong Thao

Lead Score - Case Study

Build a logistic regression model

Step 1: Reading and Understanding the Data
Step 2: Data Cleaning
Step 3: Univariate Analysis and Bi-variate Analysis
Step 4: Data Preparation
Step 5: Train-Test Split
Step 6: Feature Scaling
Step 7: Feature Selection Using RFE
Step 9: **Model Building**
Step 10: Model Evaluation
Step 11: Plotting the ROC Curve
Step 12: Finding Optimal Cutoff Point
Step 13: **Metrics - Precision and Recall**
Step 14: **Making predictions on the test set**
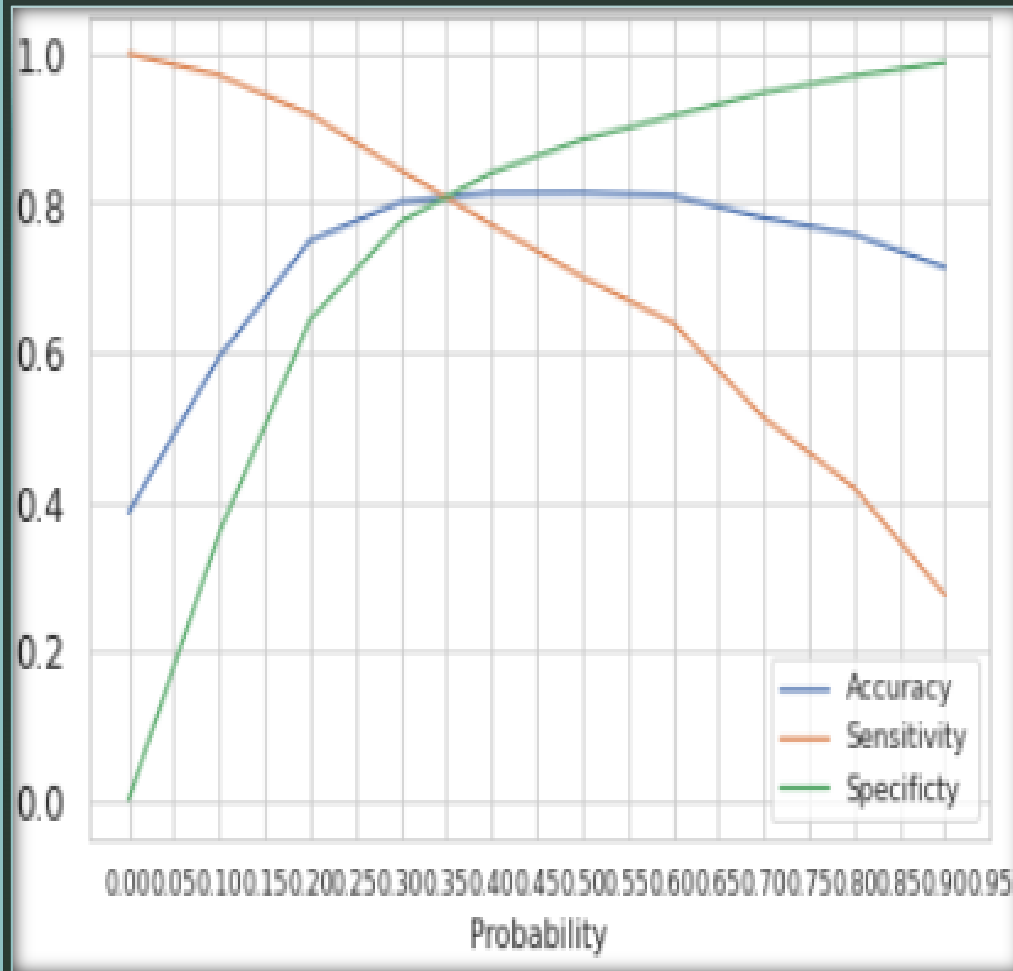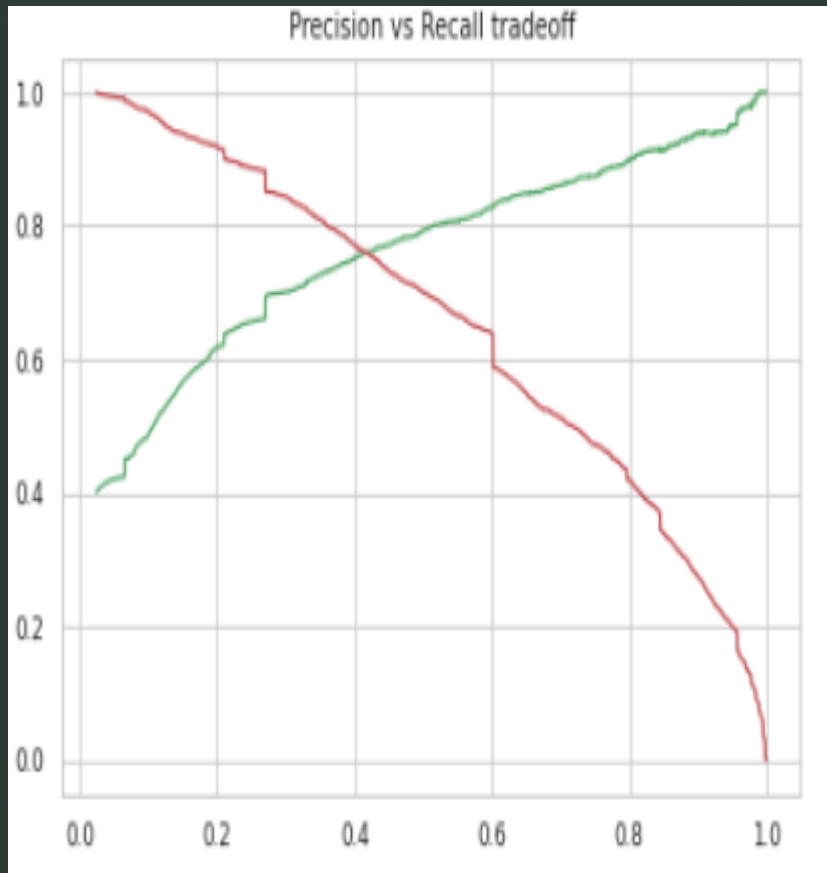
Find out problems and recommendation

- The sales team of the X-Education should focus on the leads having lead origin - lead add form , occupation - Working Professional , Lead source - Wellingak website.

- Hot Leads are identified as 'Customers having lead score above 35. Sales Team of the company should first focus on the 'Hot Leads'

- The 'Cold Leads'(Customer having lead score <= 35) should be focused after the Sales Team is done with the 'Hot Leads'.

Find out problems and recommendation

- There are many important variables like city, specialization , occupation which can potentially explain Conversion better. It is important for the management to make few of these information mandatory to fill , so that we can use in our model and build important decisions for the business.

- .

Precision vs Recall tradeoff

Find out problems and recommendation

- We have high recall score than precision score. Hence this model has an ability to adjust with the company's requirements in coming future.

Find out problems and recommendation

- High Sensitivity will ensure that almost all leads who are likely to Convert are correctly predicted where as high Specificity will ensure that leads that are on the brink of the probability of getting Converted or not are not selected

- Thank you for attention!