

Upgrad Case Study

Lead Score Case study

Group Members

- Santosh Kanmuse
- Pothurala Santhosh
- Sanjay Kale

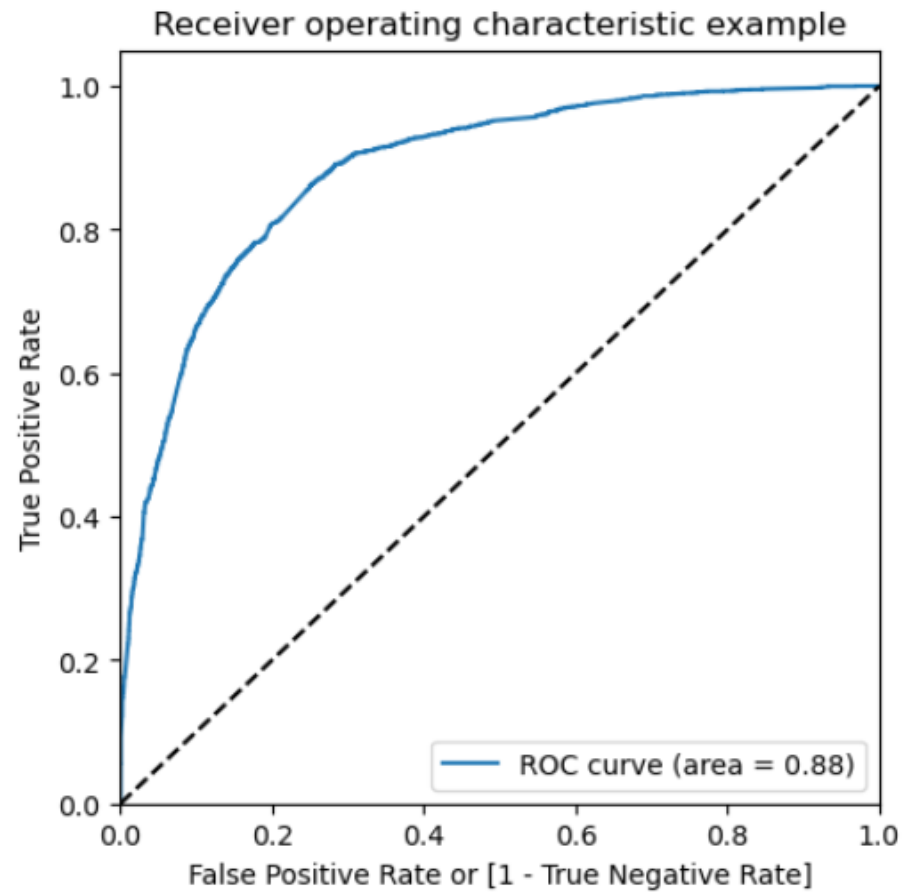
Problem Statement:

Education institute sells professional courses to professionals. It gets lot of leads through various channels, but the conversion rate of leads is poor. To improve the lead conversion efficiency institute wants to identify the most potential leads using various features contributing to the lead conversion. So that sales team can focus more on these 'hot leads.

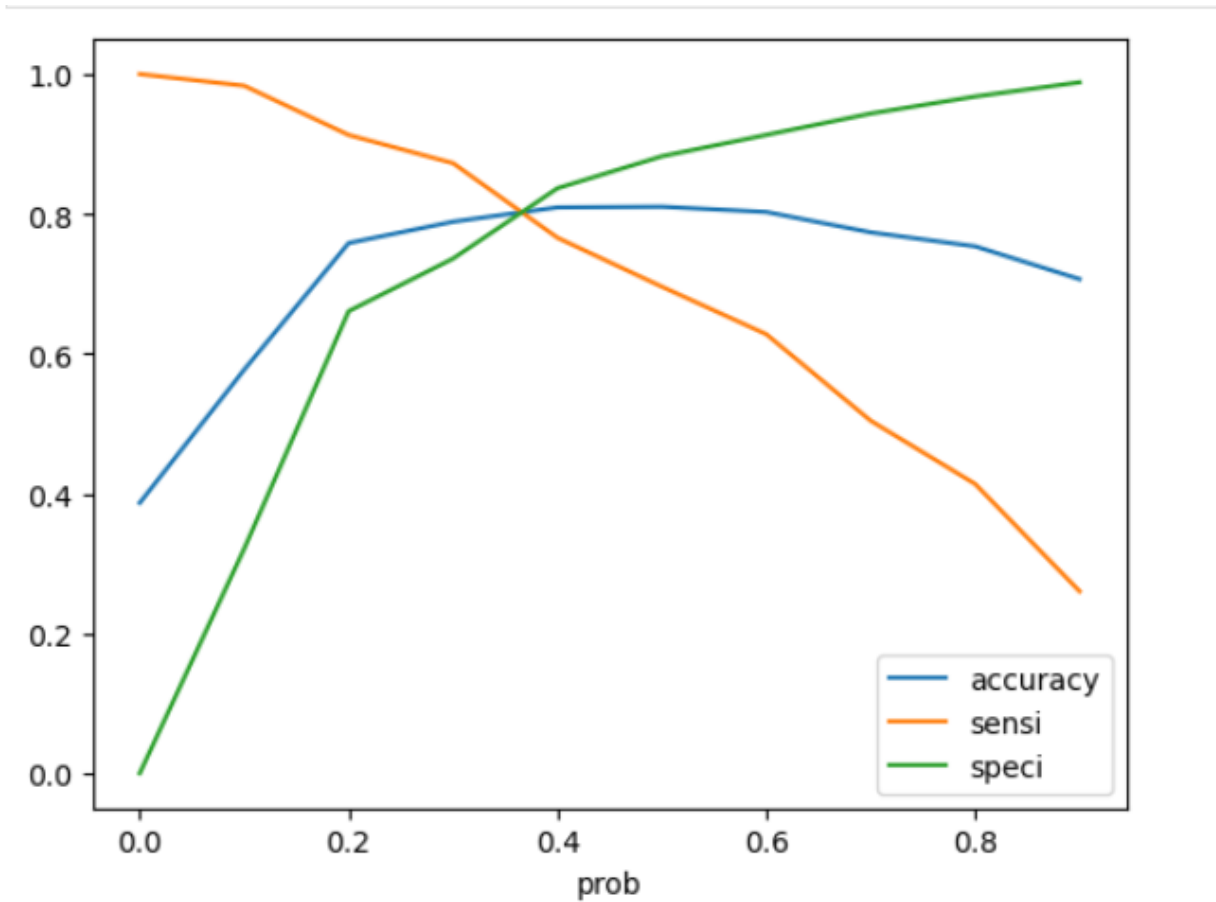
Steps followed

- Imported dataset and cleaned data which involves removing null values and adding values for missing values if needed.
- Performed EDA to understand the various trends and relations between features.
- Feature scaling using dummy variables.
- Built model using logistic regression and eliminated features by considering VIF and p values.
- Finalized the model and derived the optimum cutoff and achieved the accuracy, sensitivity and specificity of around 80%.

ROC curve



Optimum cutoff



Conclusion

It was found that the variables that mattered the most in the potential buyers are (In descending order) :

- The total time spend on the Website.
- Total number of visits
- When the last activity was:
 1. SMS
 2. Olark chat conversion
- When the google source was:
 1. Google
 2. Direct Traffic
 3. Organic search