

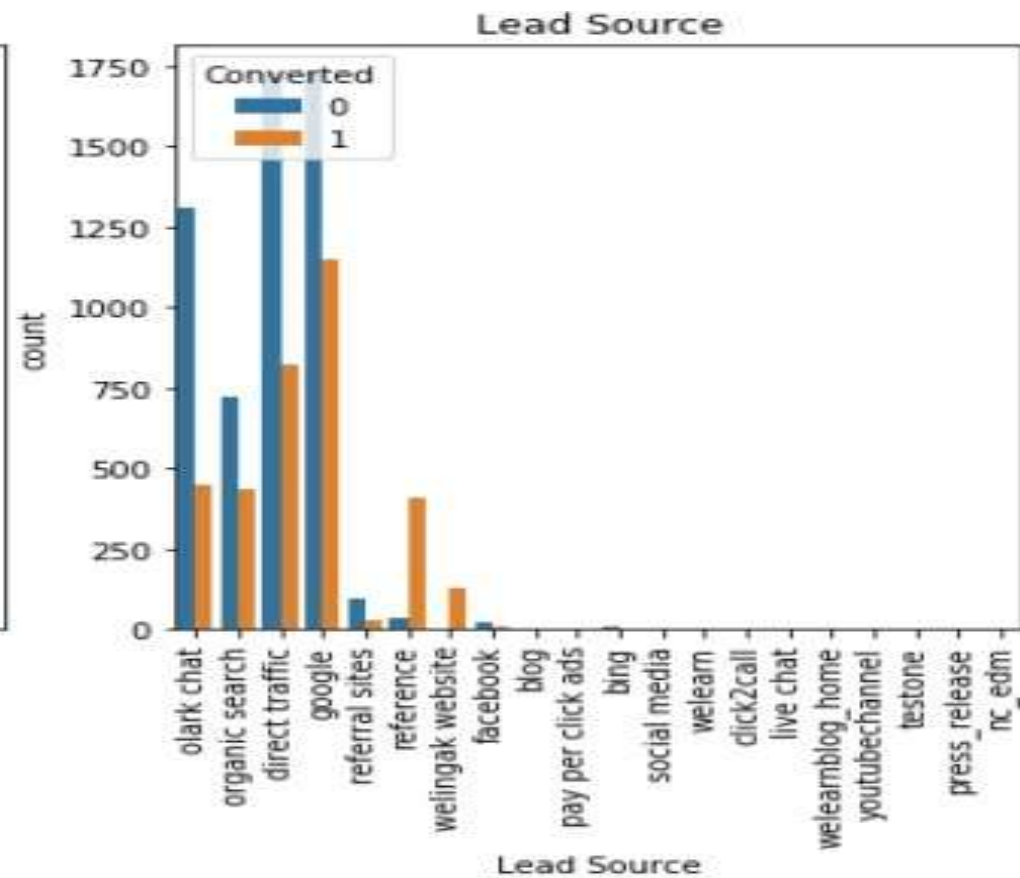
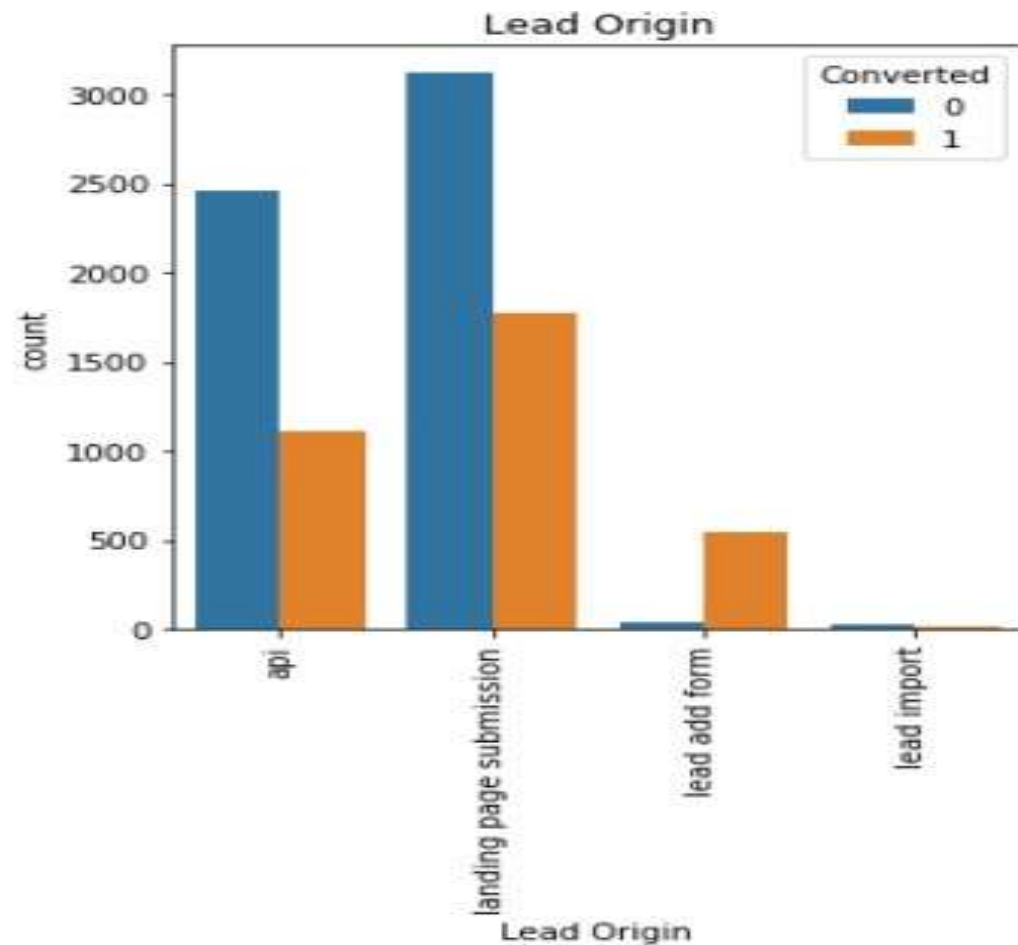
Lead Score Case Study

- Aniruddha Basuri
- Lokesh Rajendran

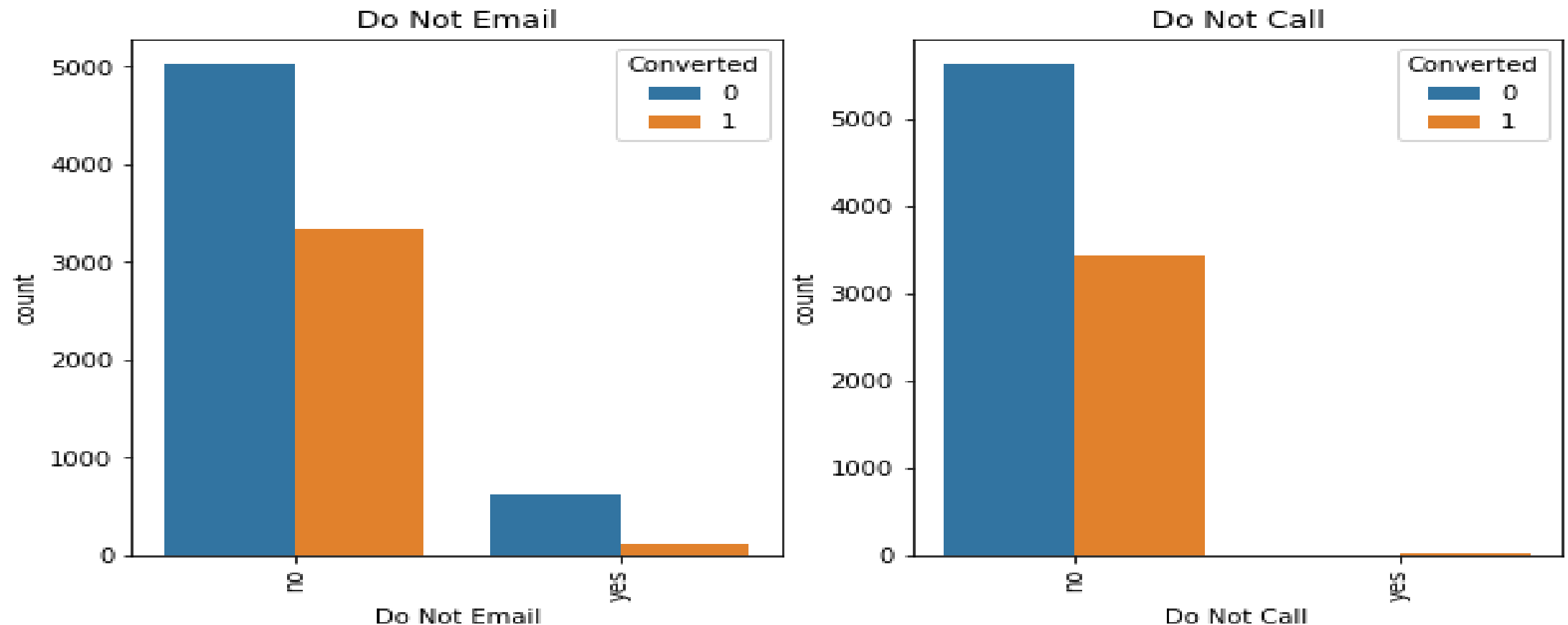
Business Problem Statement

- X Education sells online courses to industry professionals.
- X Education gets a lot of leads, its lead conversion rate is very poor. For example, if, say, they acquire 100 leads in a day, only about 30 of them are converted.
- To make this process more efficient, the company wishes to identify the most potential leads, also known as 'Hot Leads'.
- If they successfully identify this set of leads, the lead conversion rate should go up as the sales team will now be focusing more on communicating with the potential leads rather than making calls to everyone.

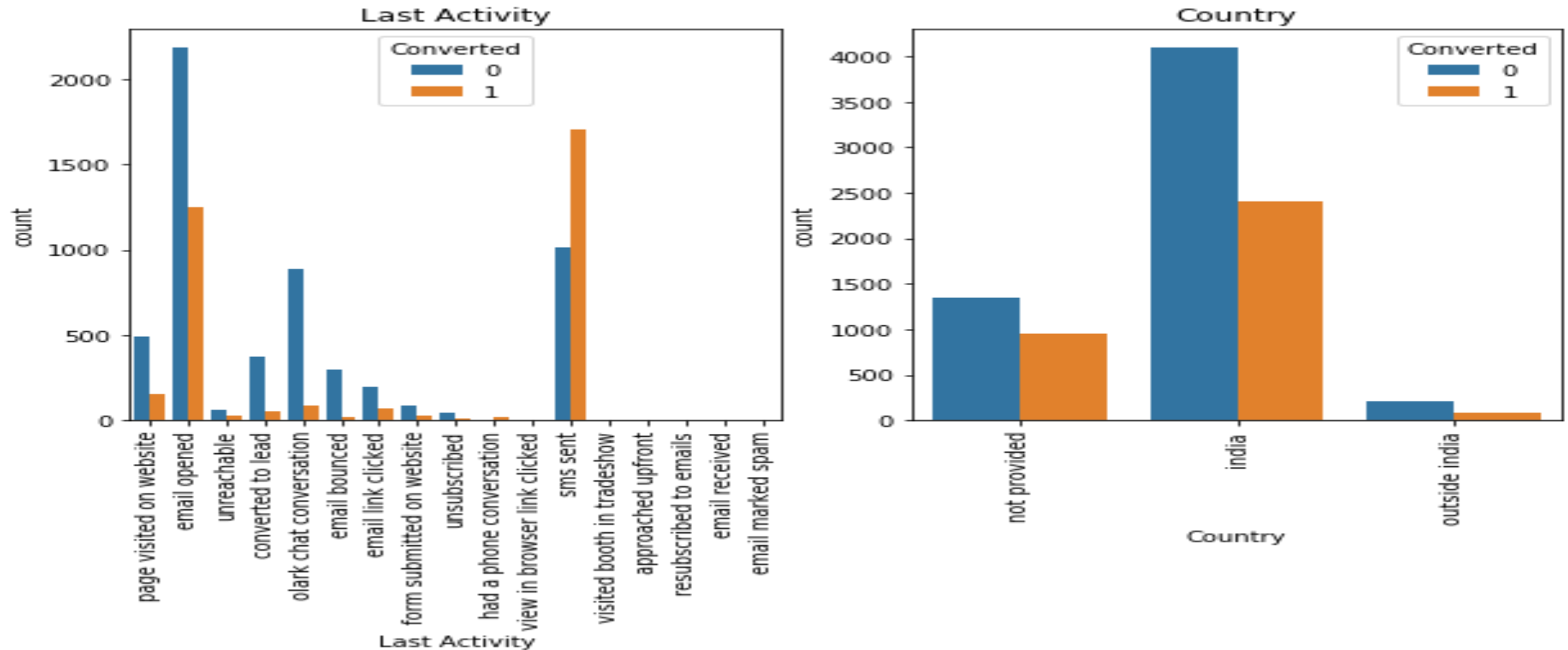
Categorical Variable Relation – Origin and Source



Categorical Variable Relation – Email and Call



Categorical Variable Relation – Activity and Country



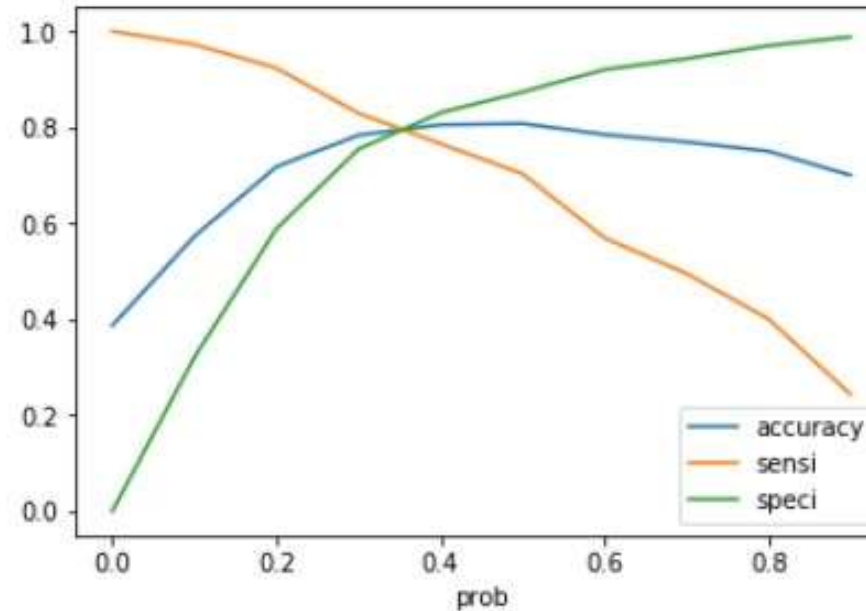
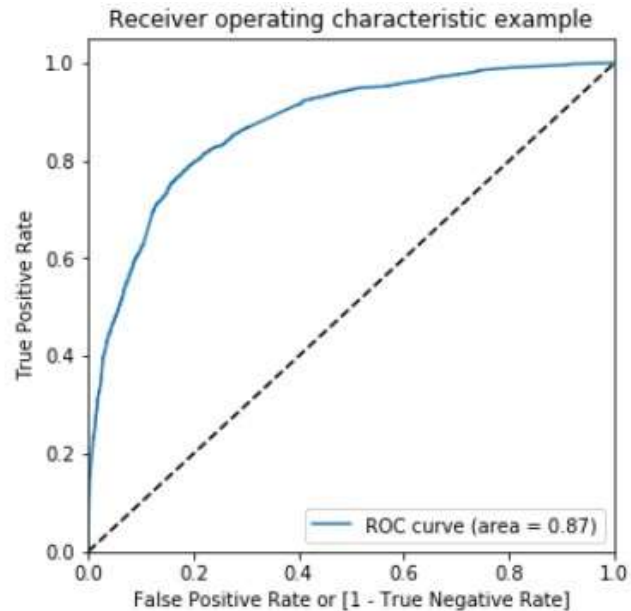
Data Normalization

- Numerical Variables are Normalized
- Dummy Variables are created for object type variables
- Total Rows for Analysis: 8792
- Total Columns for Analysis: 43

Model Building

- Splitting the Data into Training and Testing Sets
- The first basic step for regression is performing a train-test split, we have chosen 70:30 ratio.
- Use RFE for Feature Selection
- Running RFE with 15 variables as output
- Building Model by removing the variable whose p- value is greater than 0.05 and vif
- value is greater than 5
- Predictions on test data set
- Overall accuracy 81%

ROC Curve



- Finding Optimal Cut off Point
- Optimal cut off probability is that
- probability where we get balanced sensitivity and specificity.
- From the second graph it is visible that the optimal cut off is at 0.35.

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 lead source was:
 - a. Google
 - b. Direct traffic
 - c. Organic search
 - d. Welingak website
- When the last activity was:
 - a. SMS
 - b. Olark chat conversation
- When the lead origin is Lead add format.
- When their current occupation is as a working professional.
- Keeping these in mind the X Education can flourish as they have a very high chance to get almost all the potential buyers to change their mind and buy their courses.