

Lead Score Case Study

Problem Statement:

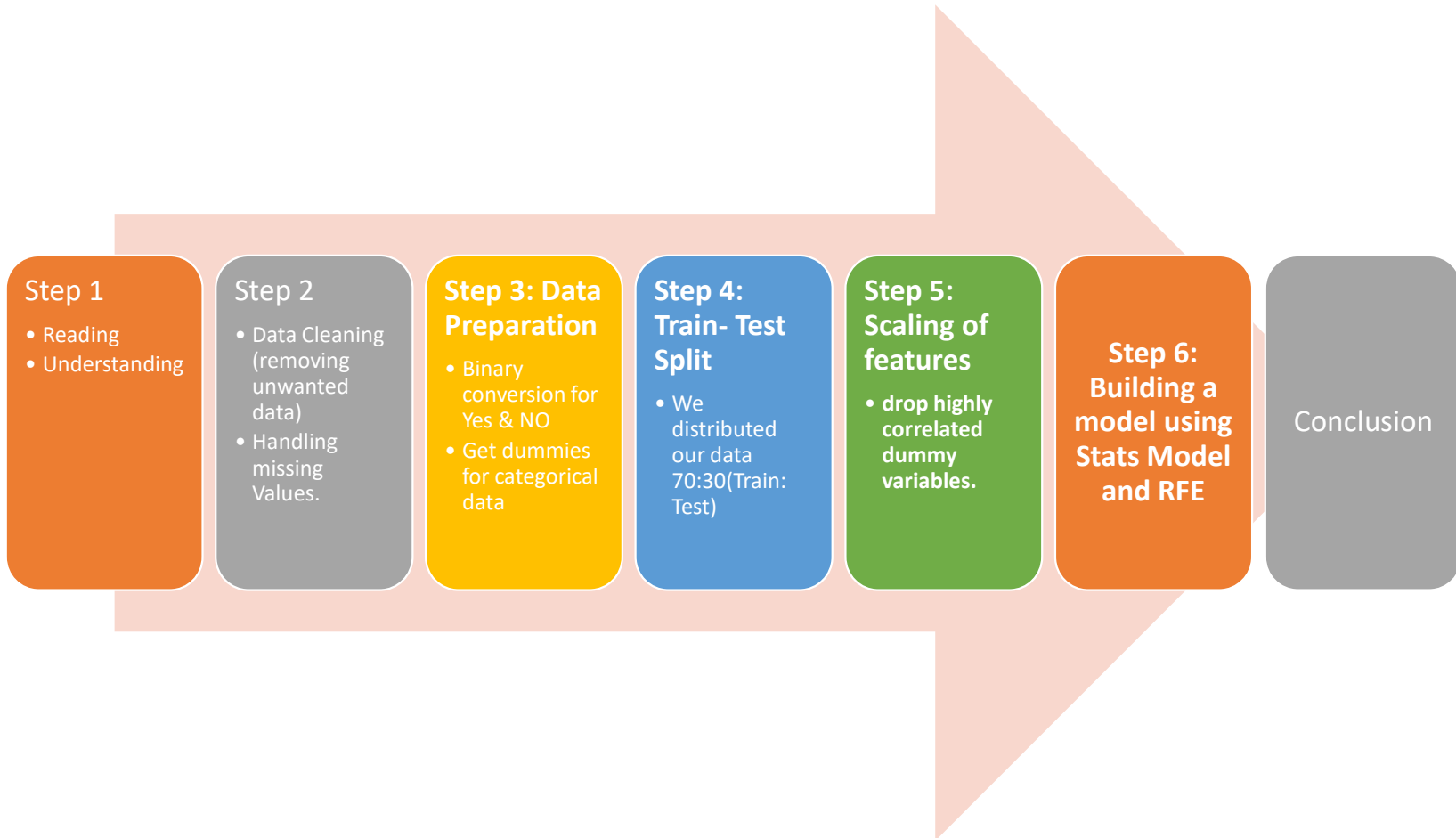
An education company named X Education sells online courses to industry professionals. On any given day, many professionals who are interested in the courses land on their website and browse for courses. For example, if, say, they acquire 100 leads in a day, only about 30 of them are converted. To make 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.

Business Goal:

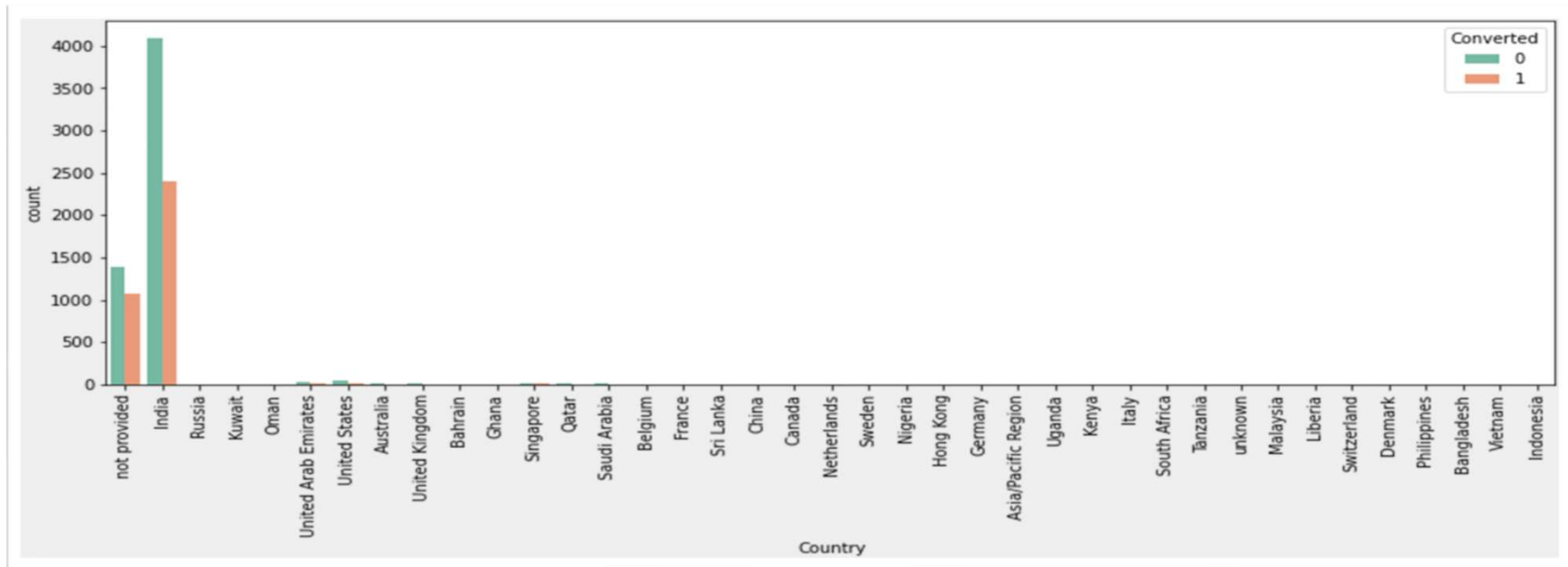
- X Education wants to select the most promising leads, i.e. the leads that are most likely to convert into paying customers.
- The company wants to build a model wherein to assign a lead score to each of the leads such that the customers with a higher lead score have a higher conversion chance and the 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%.

Lead Score Case Study

STEP TAKEN ARE AS FOLLOWS:

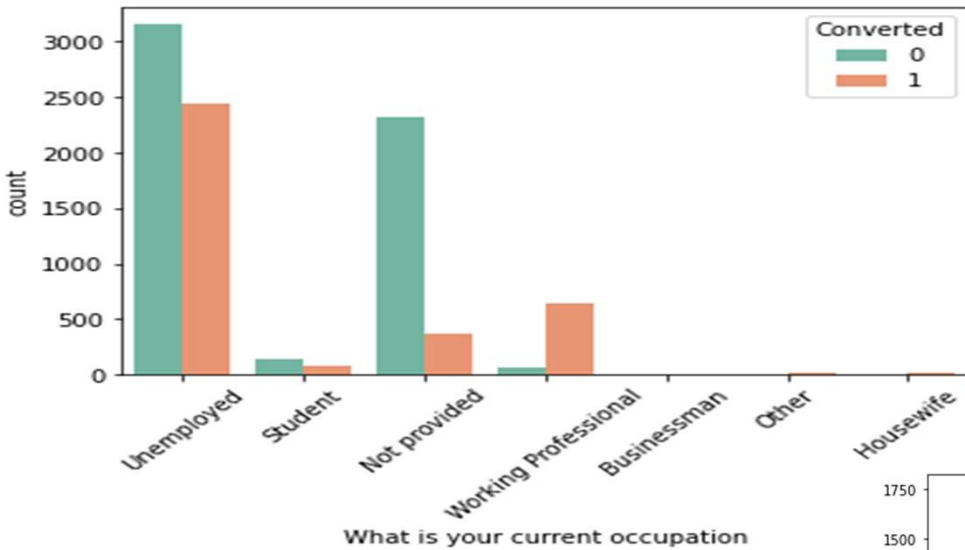


Lead Score Case Study



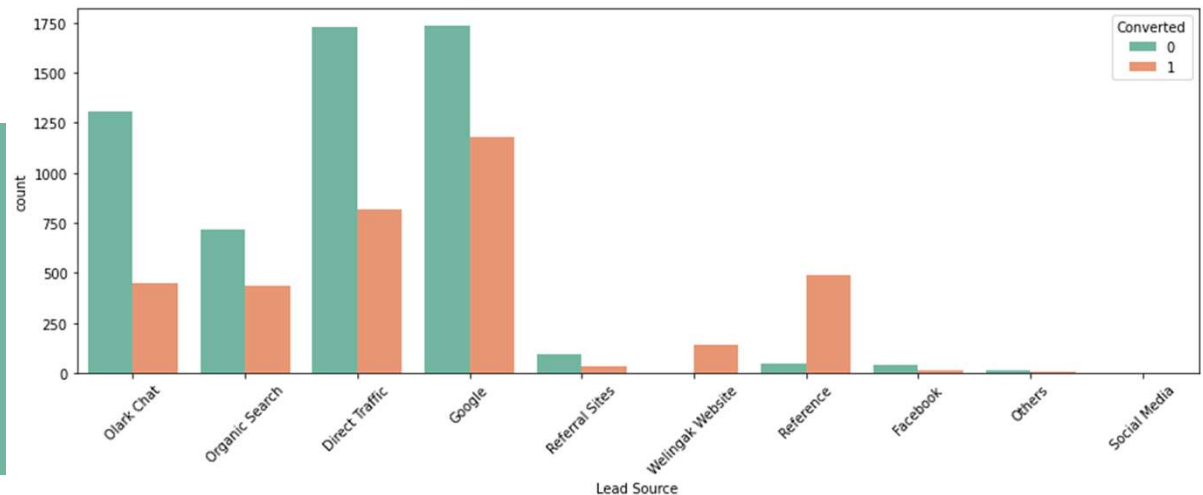
Checking for the country variable where we saw in consists of many countries data along with **India** and **Not Provided**. As we can see that most of the Leads belongs to 'India'.

Lead Score Case Study



Maximum leads generated are **unemployed** and its conversion rate is **>50%**.

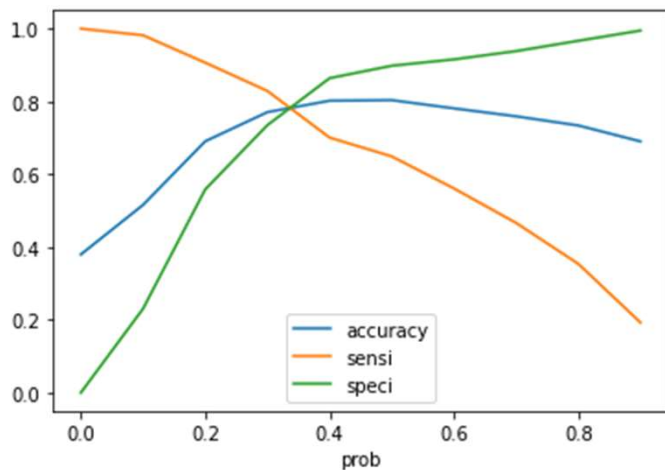
As we can see 'Google' having highest number of occurrences, hence we will impute the missing values with label 'Google'. Conversion rate of Reference leads and Welinkgak Website leads is very high. Most Leads are generated by 'Google' and 'Direct Traffic'



Lead Score Case Study

Model Evaluation - Sensitivity and Specificity on Train Data Set

From the above curve, 0.3 is the optimum point to take it as a cutoff probability.



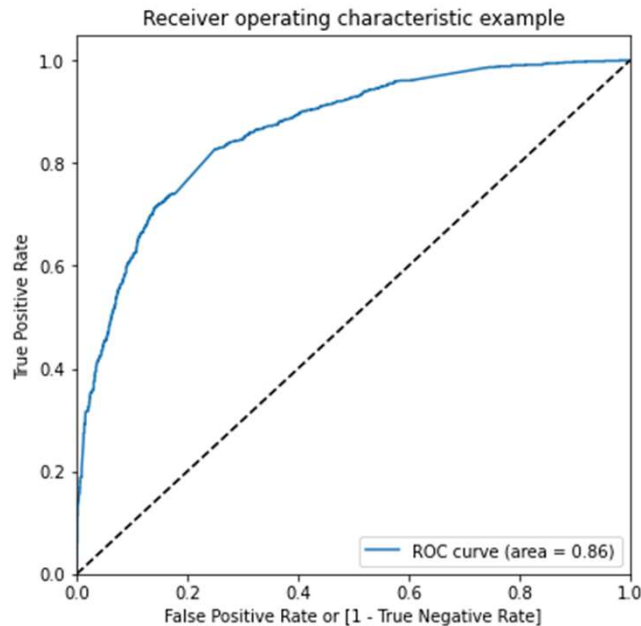
As we can see above the model is performing well. The ROC curve has a value of 0.86, which is good value. We got the following values for the Train Data:

Accuracy : 77.05%

Sensitivity : 82.89%

Specificity : 73.49%

Lead Score Case Study



As we can see above the model is performing well. The ROC curve has a value of 0.86, which is good value. We got the following values for the Test Data:

Accuracy : 77.52%

Sensitivity : 83.01%

Specificity : 74.19%

CONCLUSION:

Important features for good conversion rate or the ones' which contributes towards the probability of a lead getting converted are :

- Lead Origin_Lead Add Form
- What is your current occupation_Working Professional
- Total Time Spent on Website