LEAD SCORING CASE STUDY

SUBMITTED BY

ANUJA WALCHALE

ANIL KRISHNAN M

ANURADHA

CONTENTS

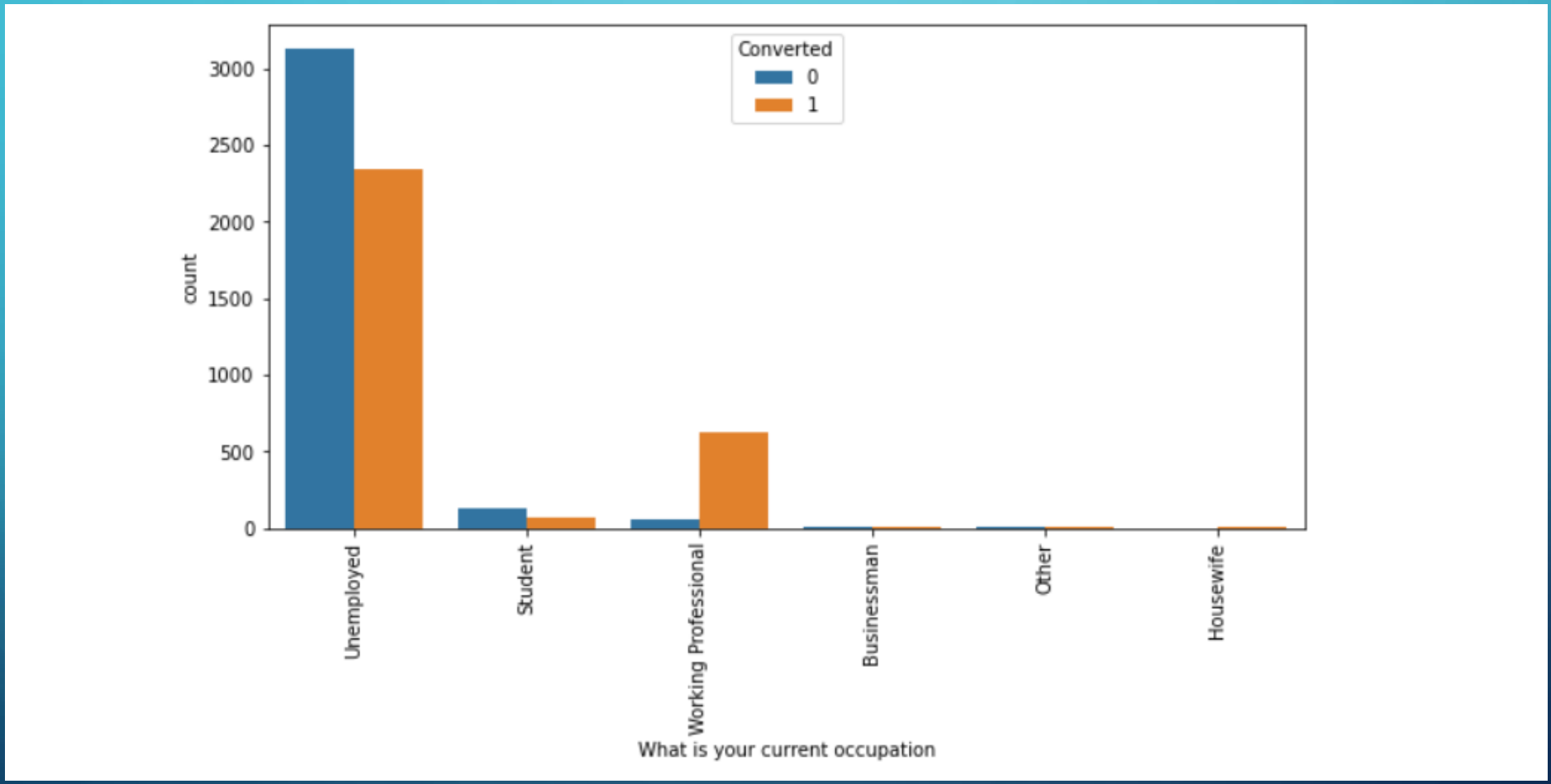
PROBLEM STATEMENT

- 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. They have process of form filling on their website after which the company that individual as a lead.
- Once these leads are acquired, employees from the sales team start making calls, writing emails, etc.Through this process, some of the leads get converted while most do not.
- The typical lead conversion rate at X education is around 30%. Now, this means 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

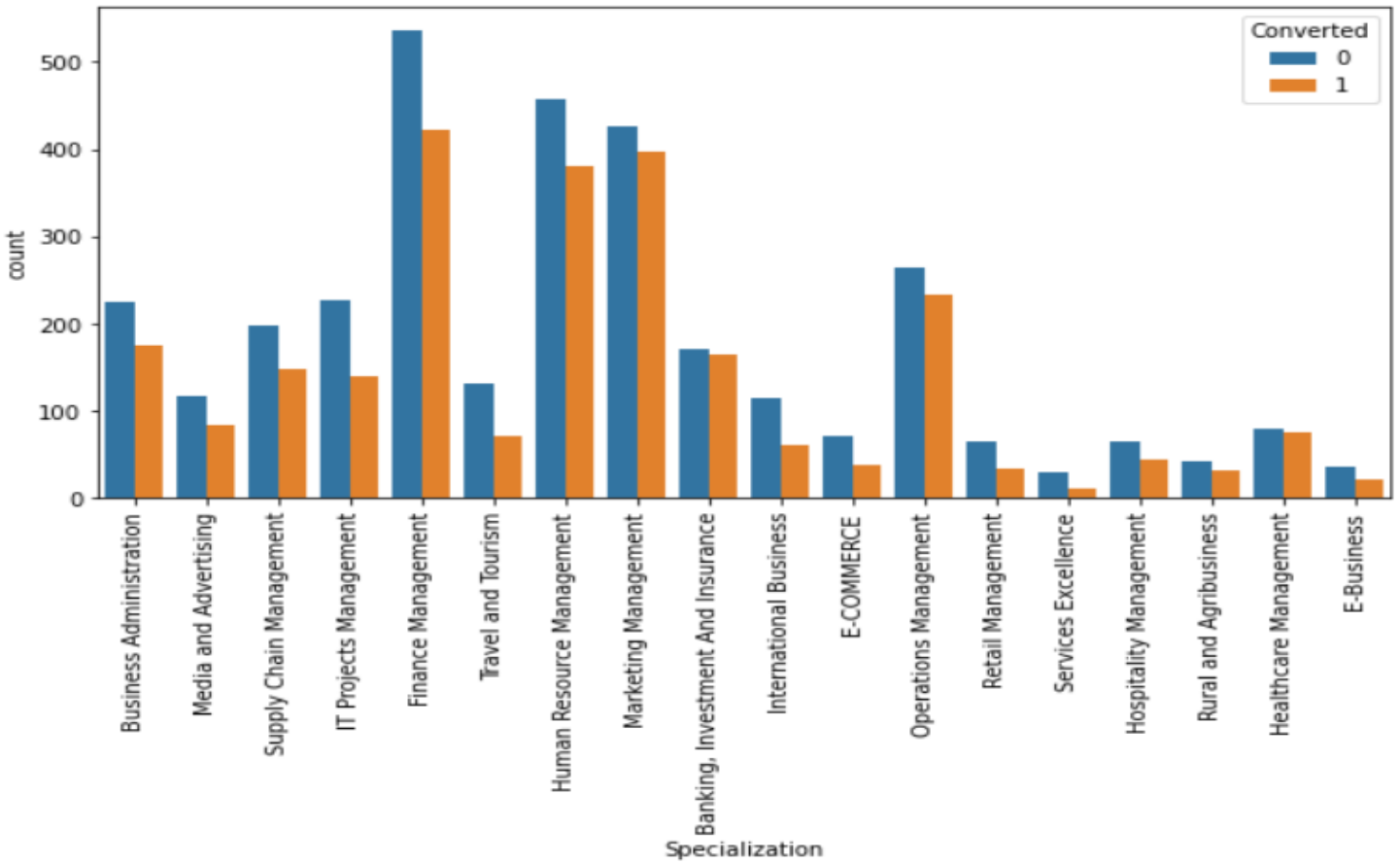
BUSINESS OBJECTIVE

- Lead X wants us to build a model to give every lead a lead score between 0 -100 . So that they can identify the Hot leads and increase their conversion rate as well.
- The CEO want to achieve a lead conversion rate of 80%.
- They want the model to be able to handle future constraints as well like Peak time actions required, how to utilize full man power and after achieving target what should be the approaches.

LEADS WHICH ARE UNEMPLOYED ARE MORE INTERESTED TO JOIN THE COURSE THAN OTHERS

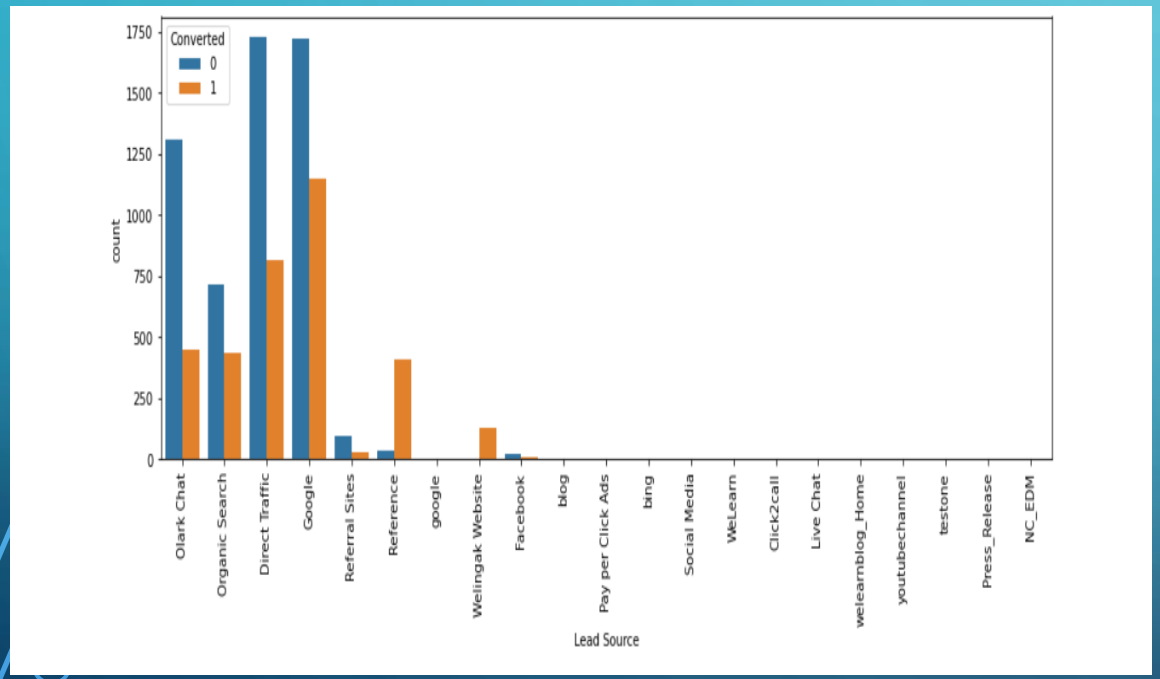


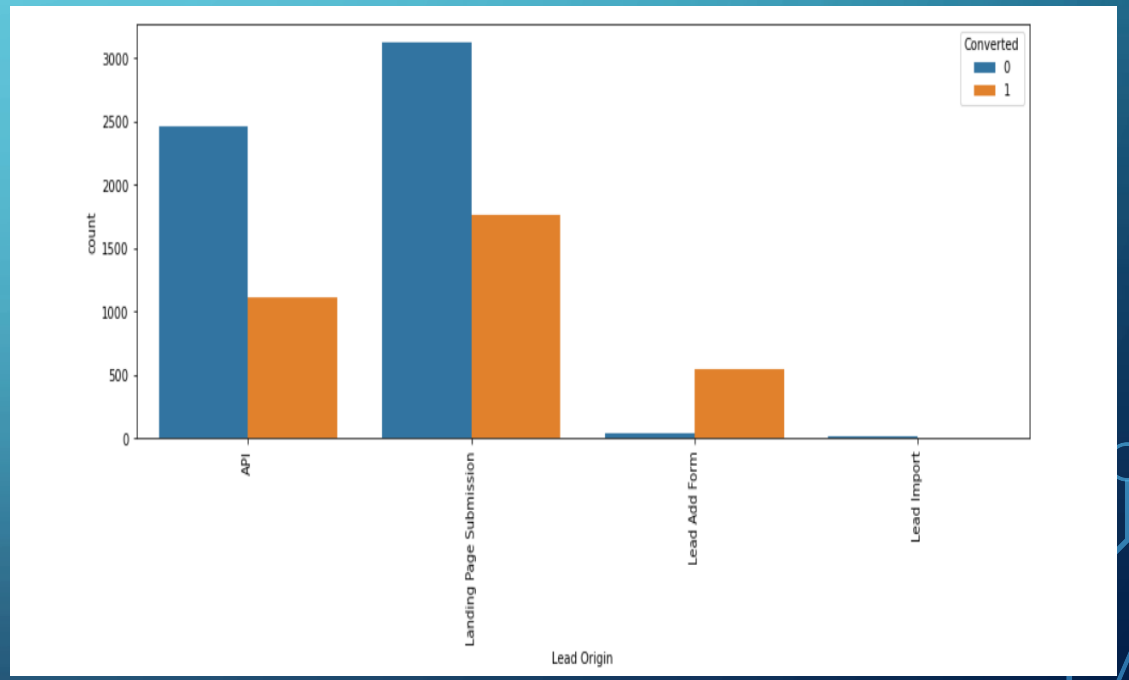
LEADS FROM HR, FINANCE & MARKETING MANAGEMENT SPECIALIZATIONS ARE HIGH PROBABILITY TO CONVERT



LEAD SOURCE & LEAD ORIGIN

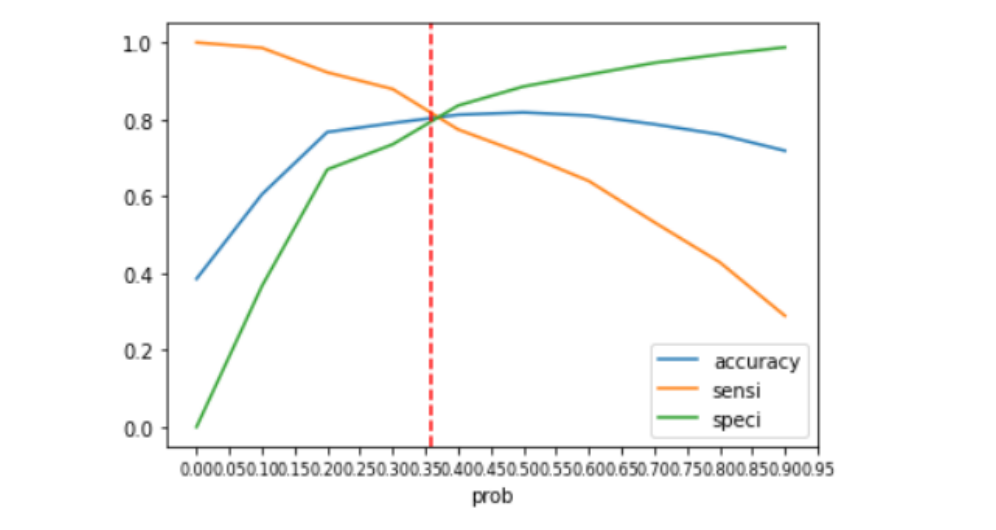
IN LEAD SOURCE THE LEADS THROUGH GOOGLE & DIRECT TRAFFIC HIGH PROBABILITY TO CONVERT WHEREAS IN LEAD ORIGIN MOST NUMBER OF LEADS ARE LANDING ON SUBMISSION

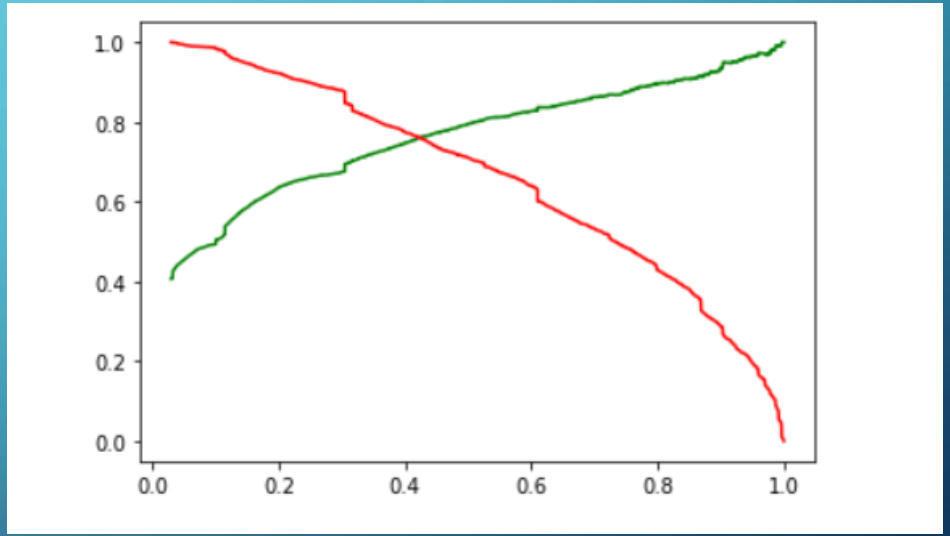




MODEL EVALUATION

0.42 is the trade-off between precision and recall Thus we can safely choose to consider any prospect lead with conversion probability higher than **42% to be a hot lead**





CONCLUSION

- 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