



Lead Score Case Study

Submitted By :
Lakshit Aswal

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.
- Now, X Education gets a lot of leads, but its lead conversion rate is very poor, only about 30%.
- 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.

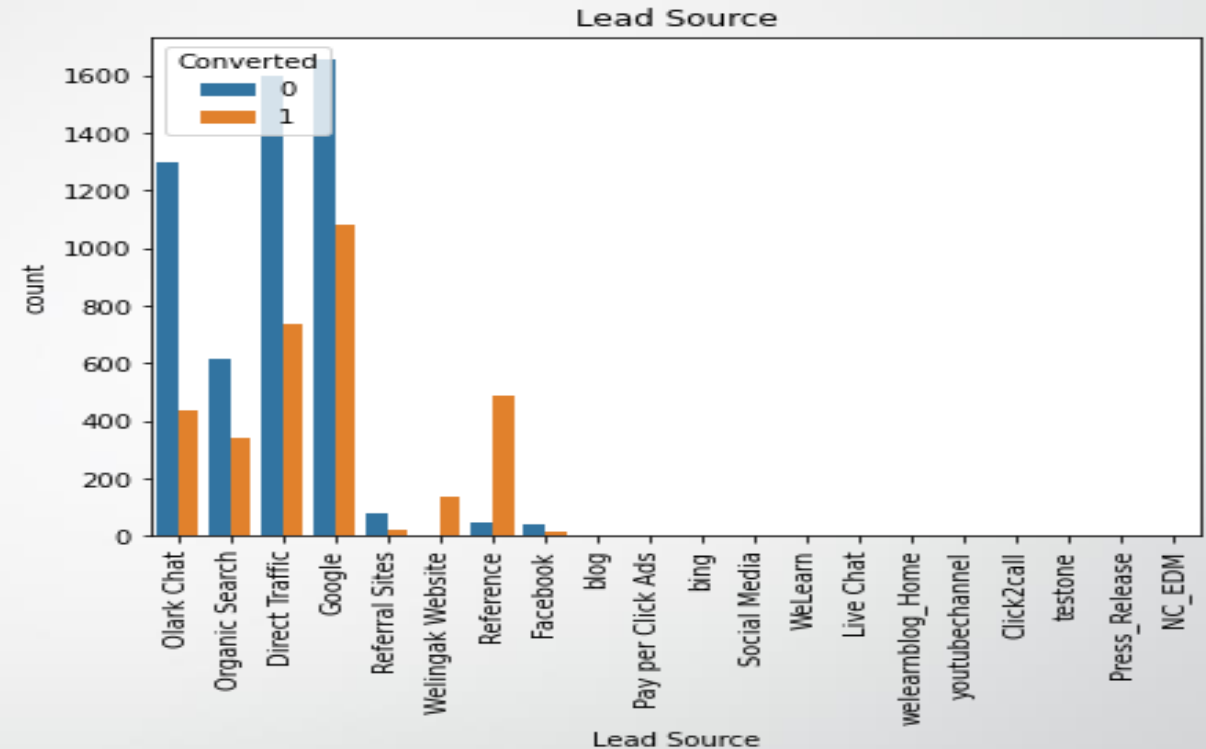
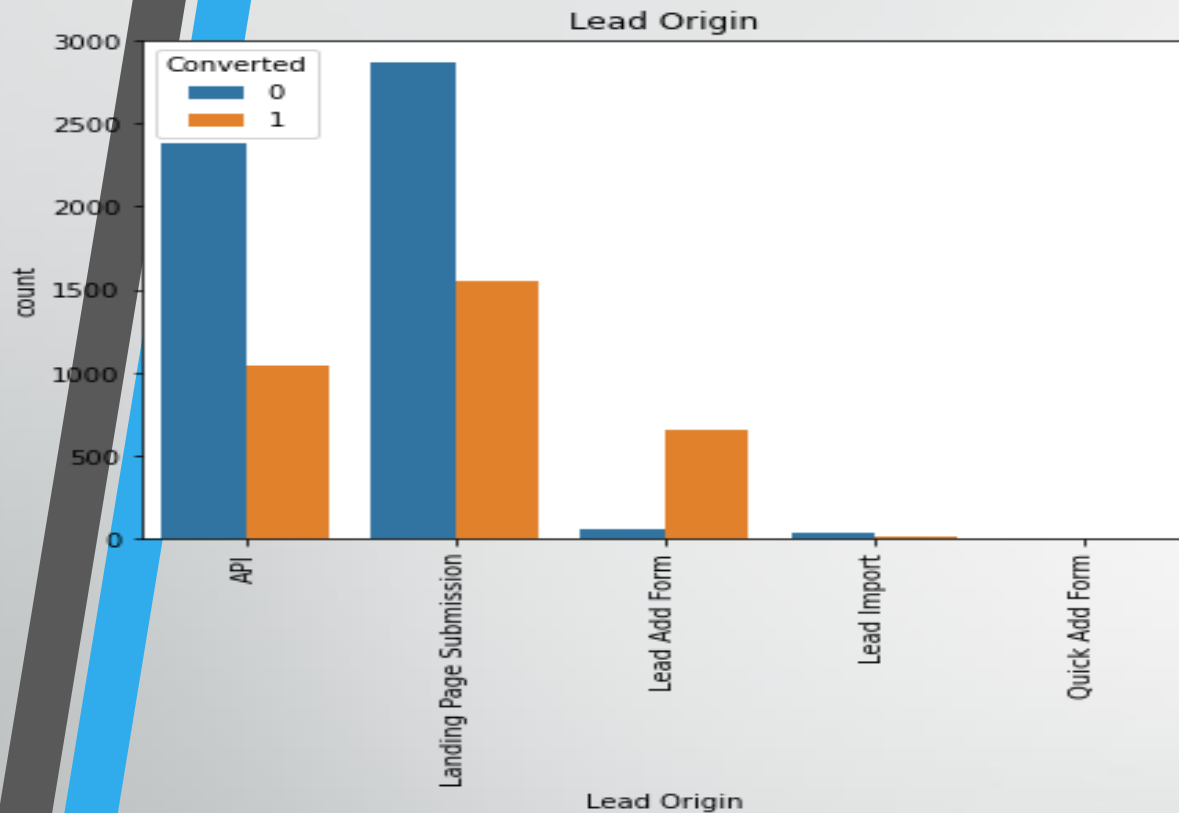
Our Goal :

- The company requires you to build a model wherein you need to assign a lead score to each of the leads such that the customers with higher lead score have a higher conversion chance and the customers with 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%.

Steps to follow :

- Extracting and understanding the dataset.
- Data cleaning and preparing.
- Exploratory data analysis to analyze the distribution and patterns.
- Feature Scaling.
- Splitting of data into train and test sets.
- Building a logistic regression model and assessing it using statsmodel.
- Evaluating the model on various metrics like Sensitivity and Specificity.
- Applying the final model on the test set and evaluating the outcome basis various metrics.

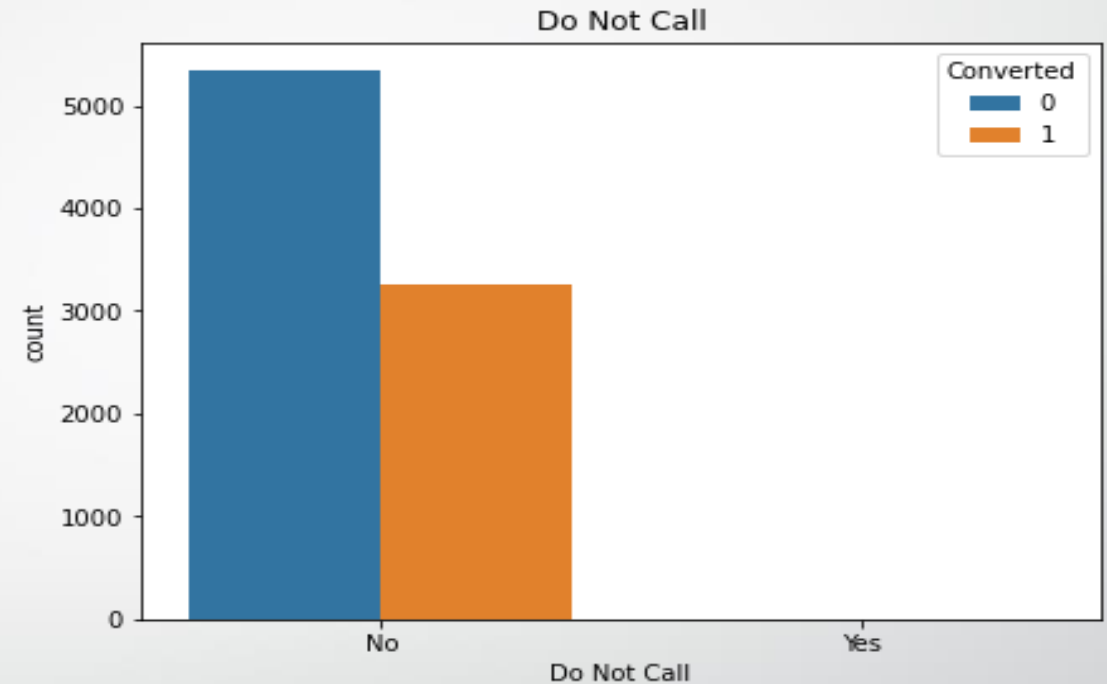
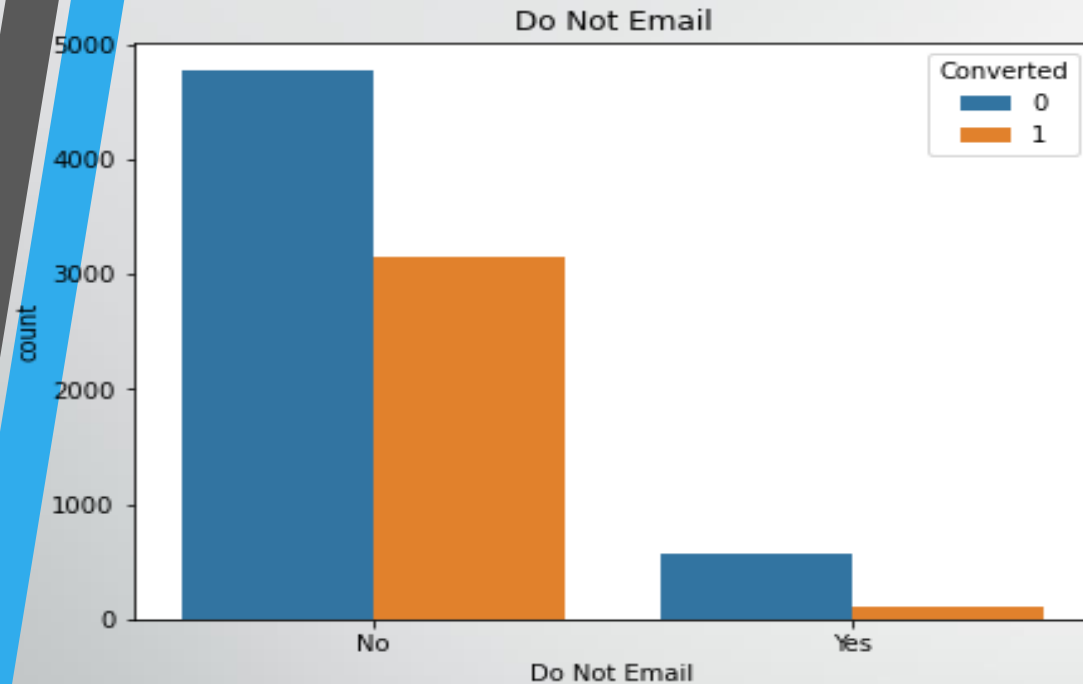
Exploratory Data Analysis : (Categorical Variables)



- Maximum conversion comes from Landing Page Submission and API.
- Also Lead Add From have better conversion rate.

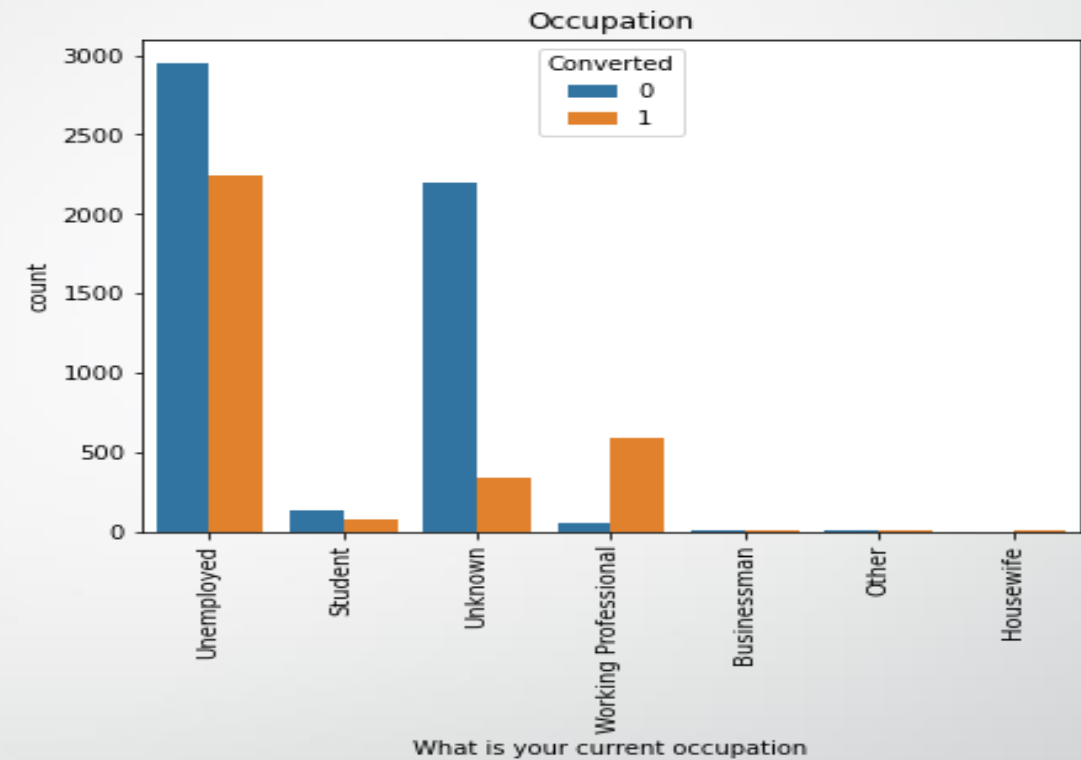
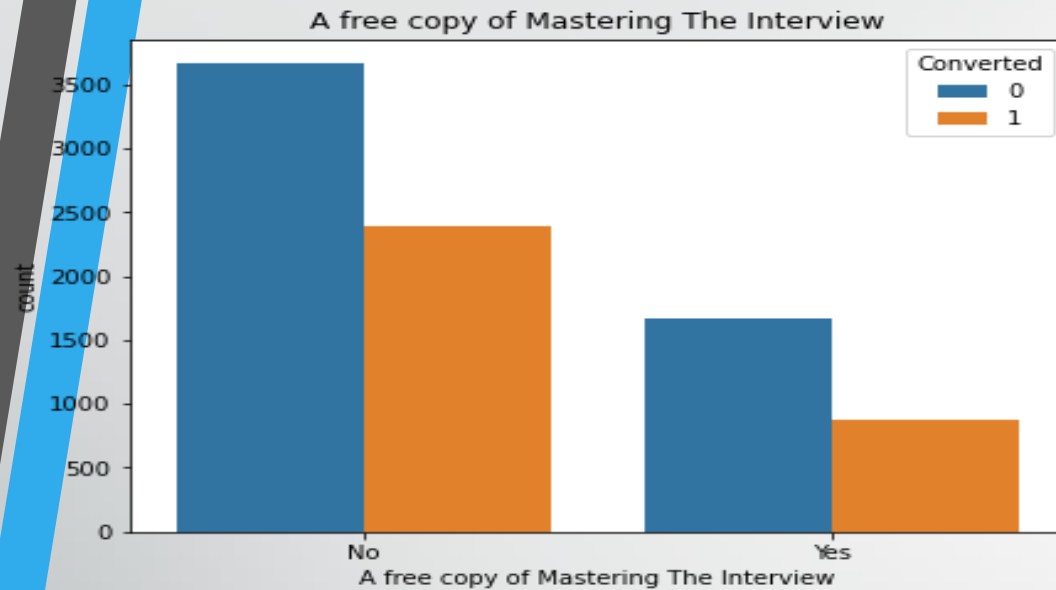
- Maximum conversion is from Google, Direct Traffic.
- Reference and Welingak Website have better conversion rate.

Exploratory Data Analysis : (Categorical Variables)



- Those opted for No Email and No Calls have more conversions.
- Very few conversions from those who opted for these service.

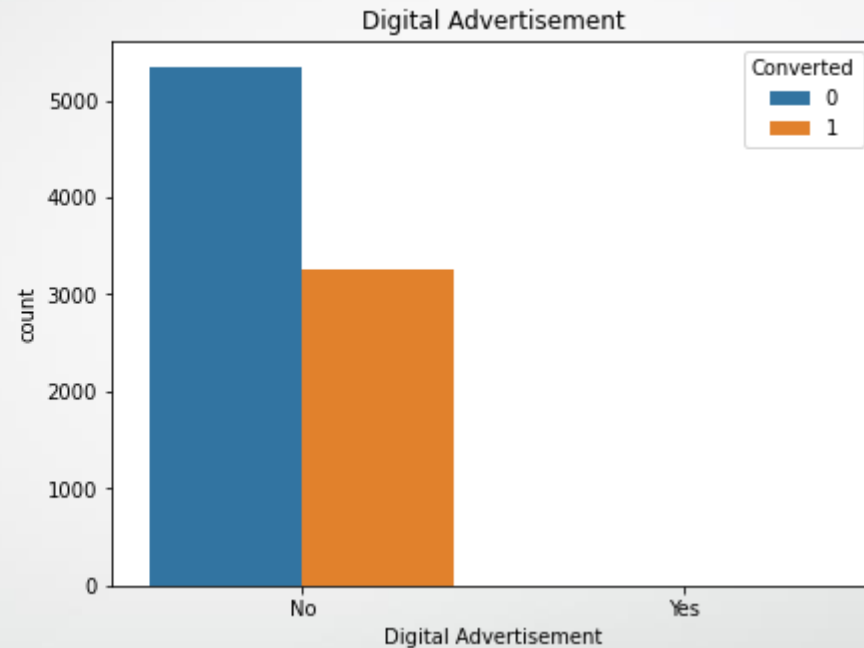
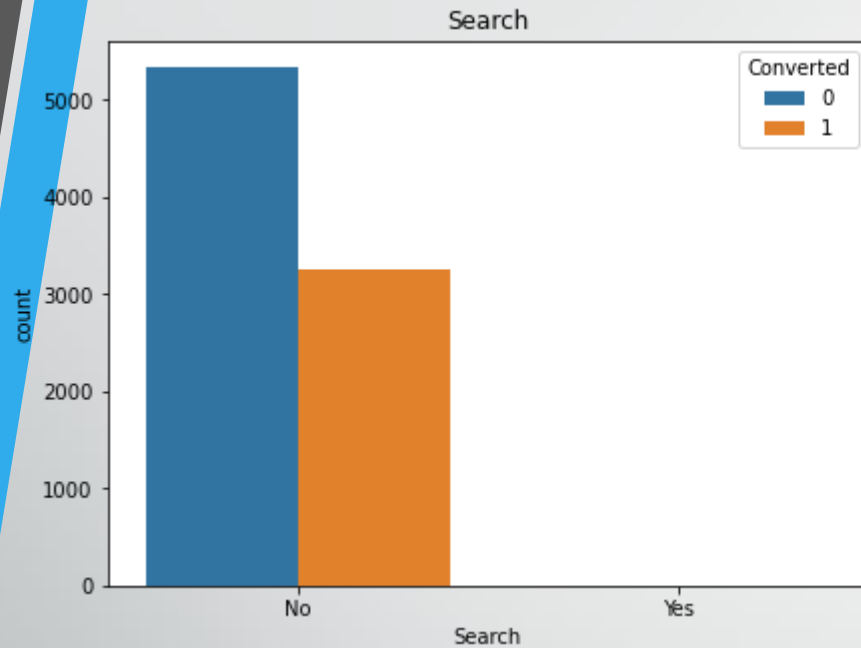
Exploratory Data Analysis : (Categorical Variables)



- Free Copy of Mastering the Interview have conversions more for those who didn't opt for it.

- More Unemployed and Working Professionals have good chance of conversion.

Exploratory Data Analysis : (Categorical Variables)



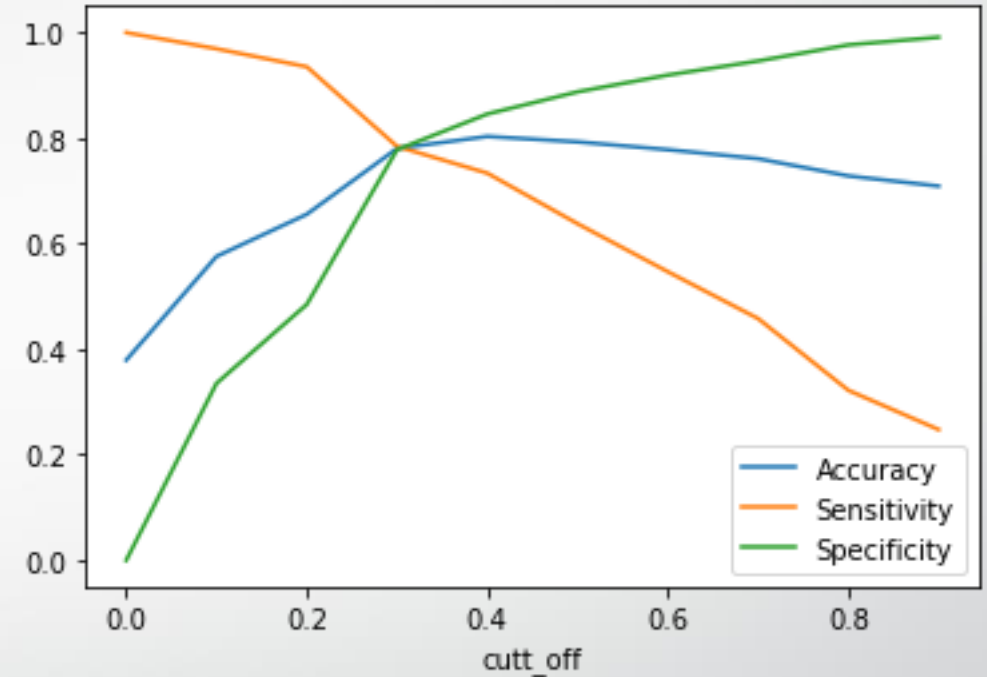
- We can't see much impact through Search and Digital Advertisement on Lead conversion.

Variables that have good impact on conversion rate:

- Do Not Email
- Total Time Spent On Website
- Lead Origin - Landing Page Submission, Lead Add Form, Lead Import
- Lead Occupation – Working Professional and Unemployed
- Lead Source – NC EDM and Welingak Website
- Why this Course - Better Career Prospects
- Country - India

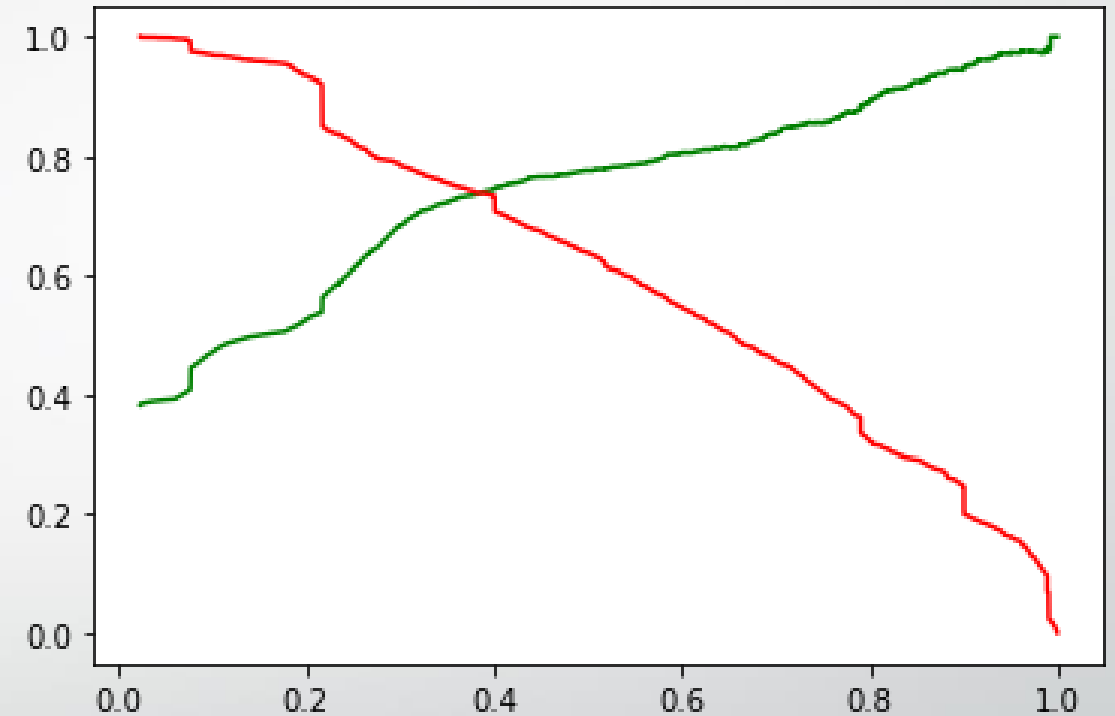
Model Evaluation – with Sensitivity and Specificity :

- The graph gave a cut-off value of 0.35.
- And below metrics were achieved.
- Accuracy – 79.9% (round to 80%)
- Sensitivity – 76.5%
- Specificity – 81.9%



Model Evaluation – with Sensitivity and Specificity :

- Precision – 72.2%
- Recall – 75.5%



Model Evaluation on test set – with Sensitivity and Specificity :

- We achieved an accuracy of 79.9% or 80% on the test set.
- The sensitivity came to be 76.5%
- The specificity came to be 81.9%

Conclusion

- The final accuracy on the test set came to be 79.9%, with sensitivity of 76% and specificity of 82%.
- Therefore the lead score we achieved through the model is around 80% which was required by our business goal.
- Also top variables to contribute on lead conversion are:
 - Total time spent on website
 - Lead origin from Lead Add Form and Landing Page Submission
 - Lead Source from Welingak Website