

X Education - Lead Scoring Case Study

Increasing lead conversion rate for X education

Group Members:

- **Ekta Davda**
- **Akash Chatterjee**
- **Chanakya Sachdeva**

Background

X Education Company

- ❖ X Education , An education company named sells online courses to industry professionals
- ❖ Many interested professionals land on their website
- ❖ The company markets its courses on several websites like Google. Once these people land on the website, they might browse the courses or fill up a form for the course or watch some videos
- ❖ When these people fill up a form providing their email address or phone number, they are classified to be 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%

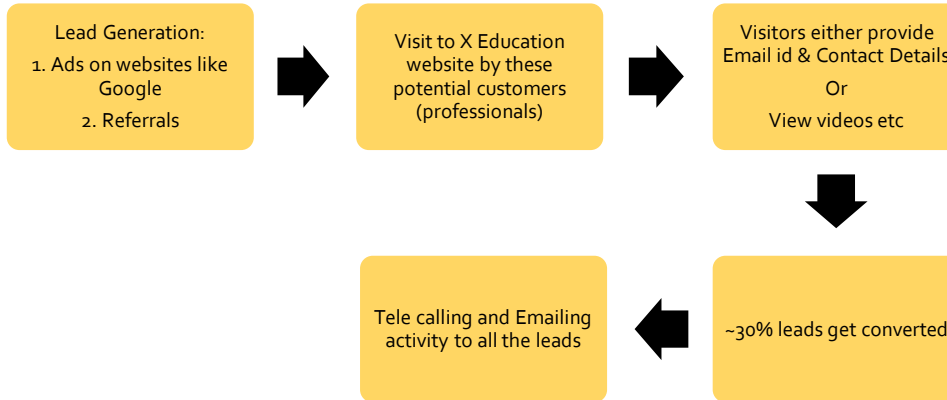
Problem Statement

X Education Company's Problem

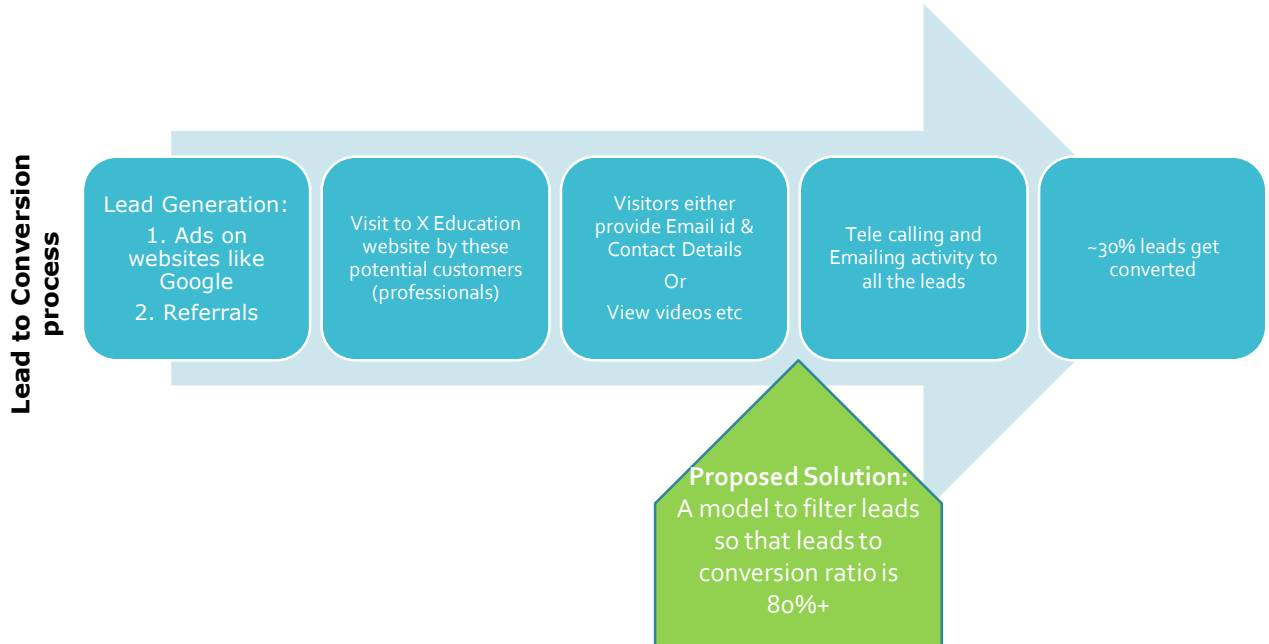
- ❖ X Education gets a lot of leads but its lead conversion rate is very poor
- ❖ To make this process more efficient, the company wishes to identify the most potential leads, also known as 'Hot Leads'
- ❖ If they successfully identify this set of leads, the lead conversion rate should go up as the sales team will now be focusing more on communicating with the potential leads rather than making calls to everyone
- ❖ We will help them to select the most promising leads, i.e. the leads that are most likely to convert into paying customers.
- ❖ We are required to build a model wherein we need to assign a lead score to each of the leads such that the customers with higher lead score have a higher conversion chance
- ❖ The CEO, in particular, has given a ballpark of the target lead conversion rate to be 80%.

Lead – Conversion Process

Lead to Conversion process



Lead – Conversion Process



Solution

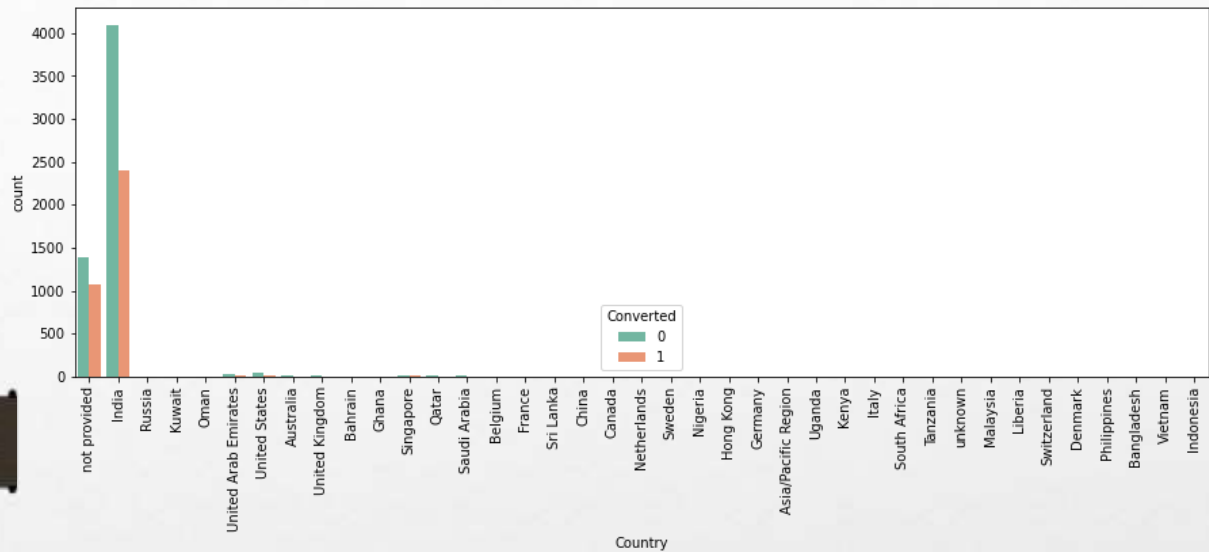
Selection of Hot Leads

For our Problem Solution, the crucial part is to accurately identify hot leads.

The more accurate we obtain the hot lead, the more chance we get of higher conversion ratio.

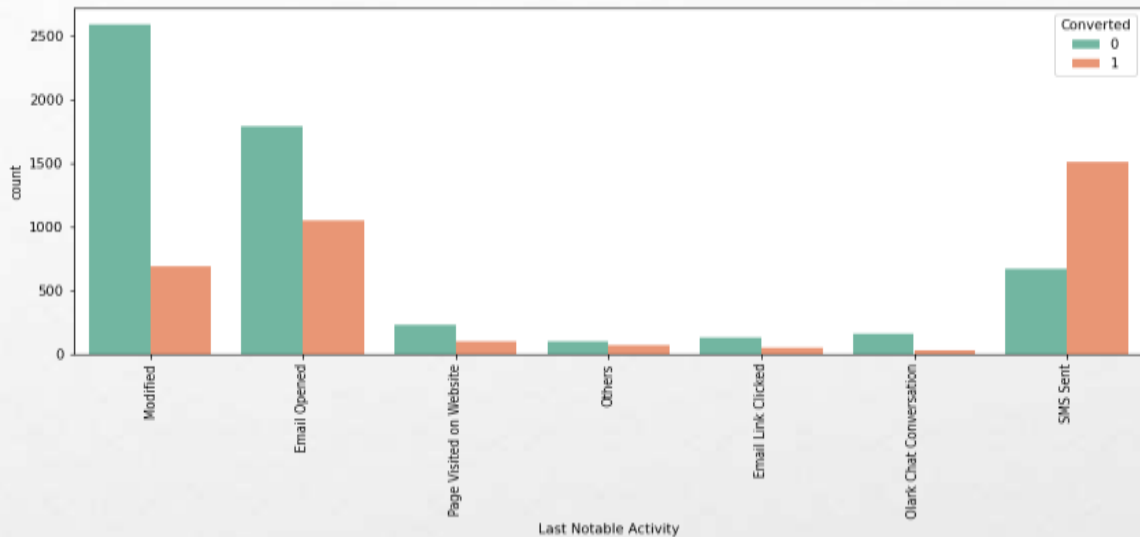
Since we have a target of 80% conversion rate, we would want to obtain a high accuracy in obtaining hot leads.

EDA- Visualization



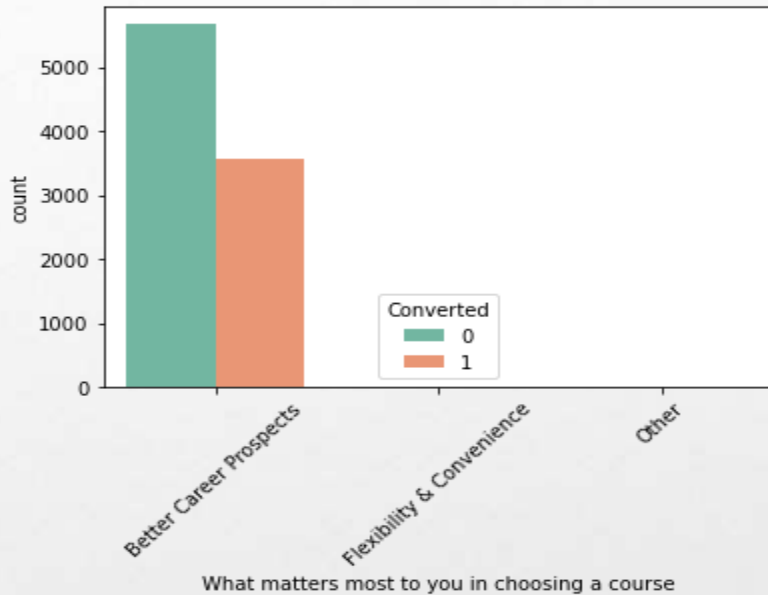
Visualizing Country variable after imputation

India is the top source of student and target working professional or unemployment.

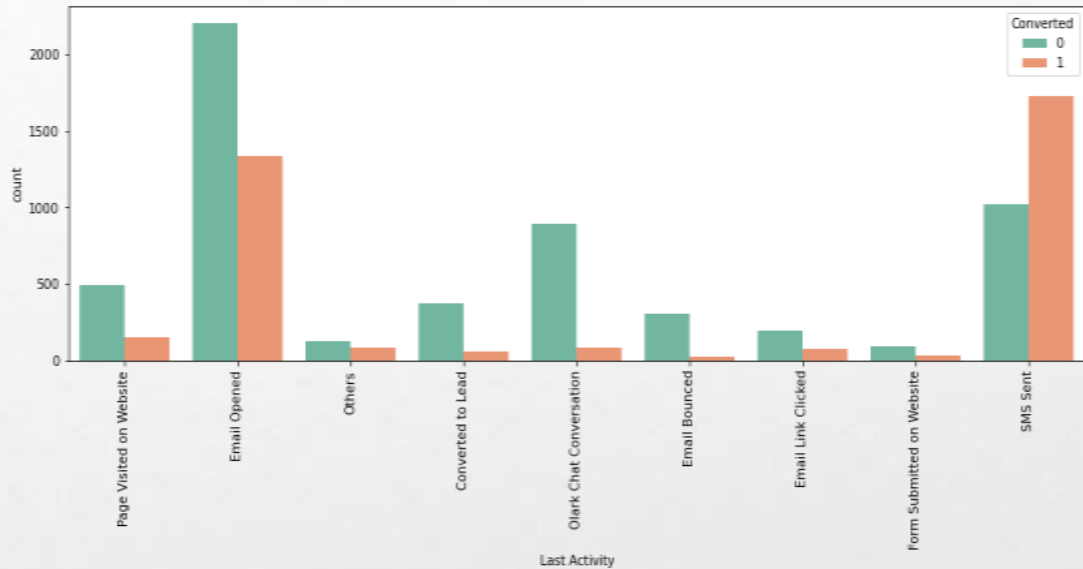


visualizing count of Variable based on Converted value

Maximum leads generated are unemployed and their conversion rate is more than 50%



visualizing count of Variable based on Converted value

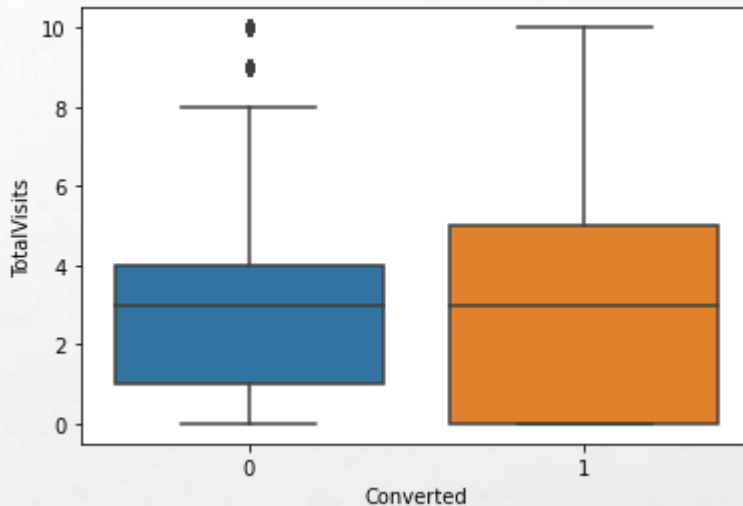


Visualizing count of Last Activity Variable

Maximum leads are generated having last activity as Email opened but conversion rate is not too good.

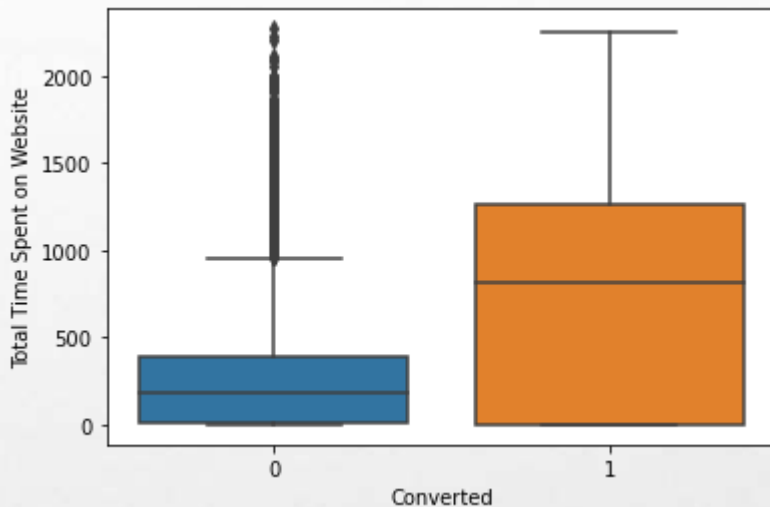


Checking correlations of numeric values using heatmap



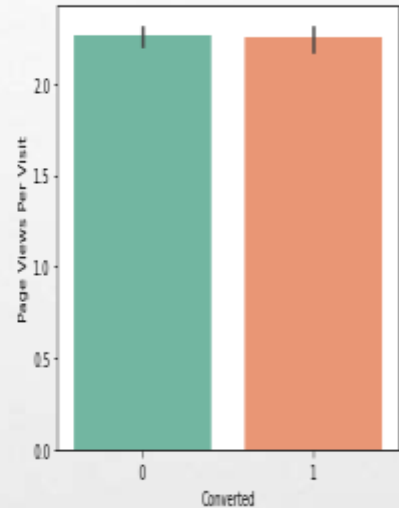
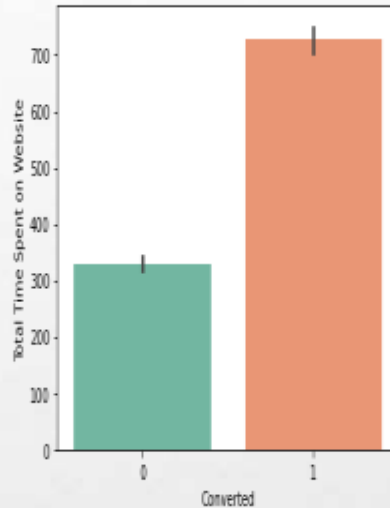
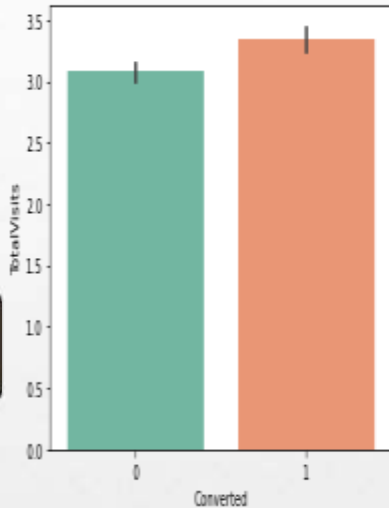
Visualizing TotalVisits w.r.t Target Variable 'Converted'

As the median for both converted and non-converted leads are same , nothing conclusive can be said on the basis of variable TotalVisits



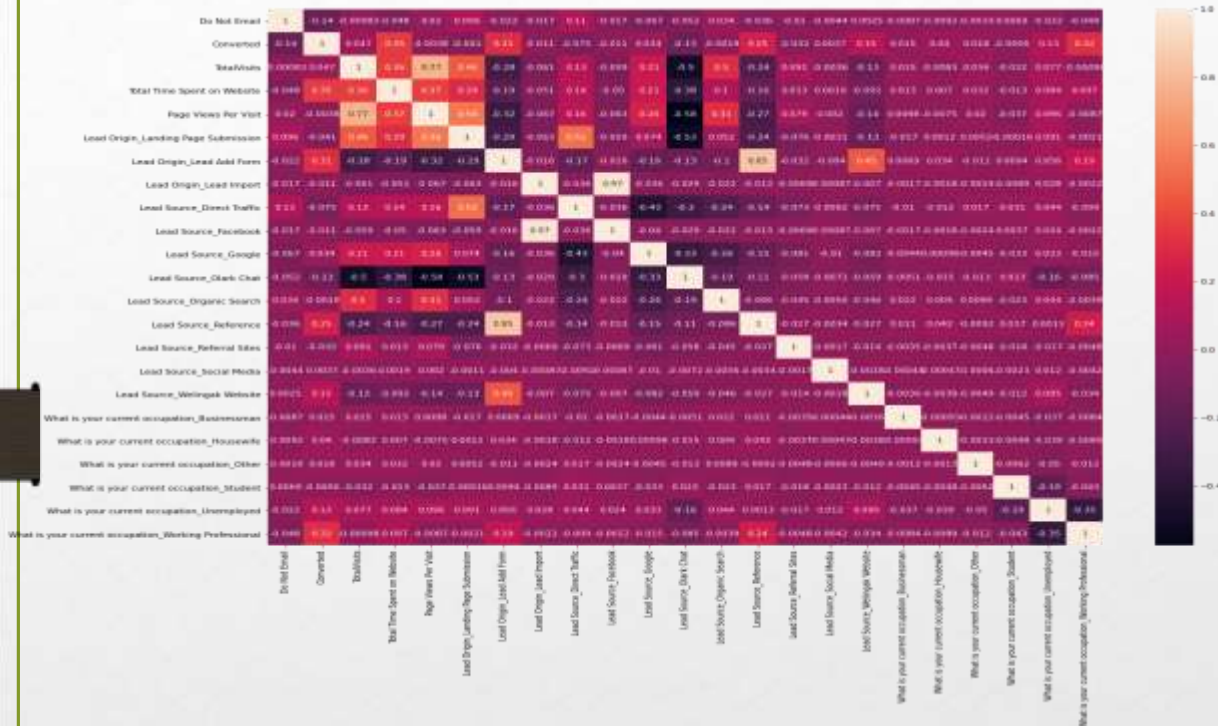
Visualizing TotalVisits w.r.t Target Variable 'Converted'

As can be seen, leads spending more time on website are more likely to convert , thus website should be made more engaging to increase conversion rate

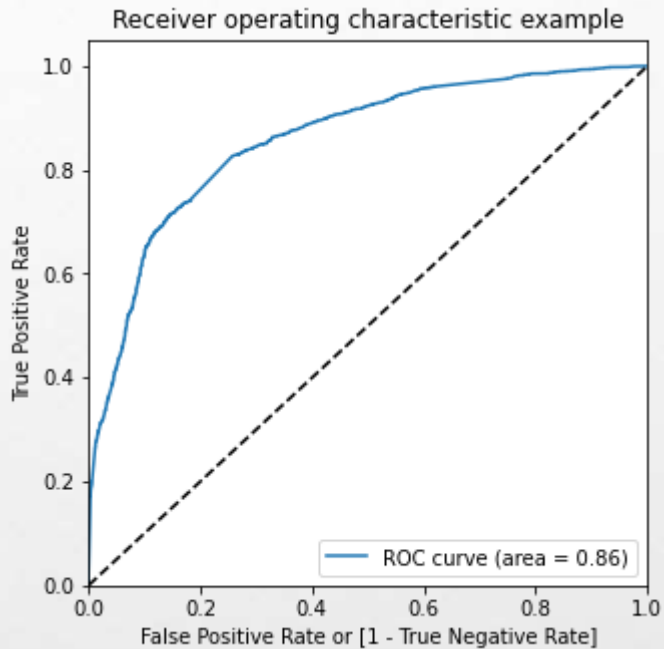


Checking the conversions for all numeric values

The conversion rate is high for Total Visits, Total Time Spent on Website and Page Views Per Visit

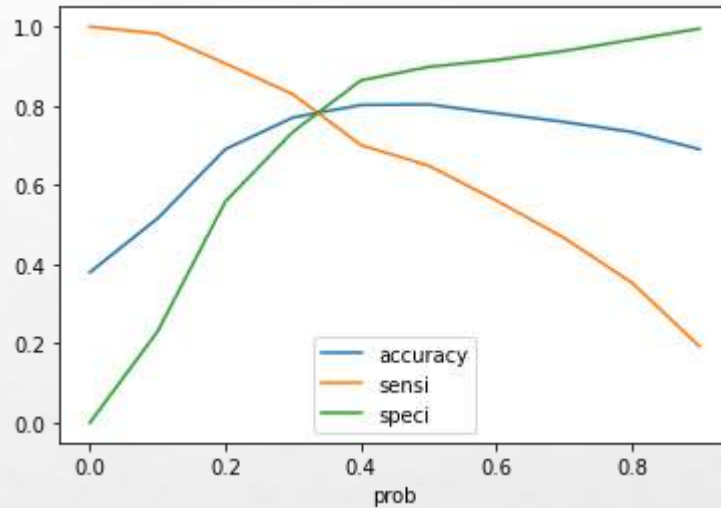


Correlation Matrix



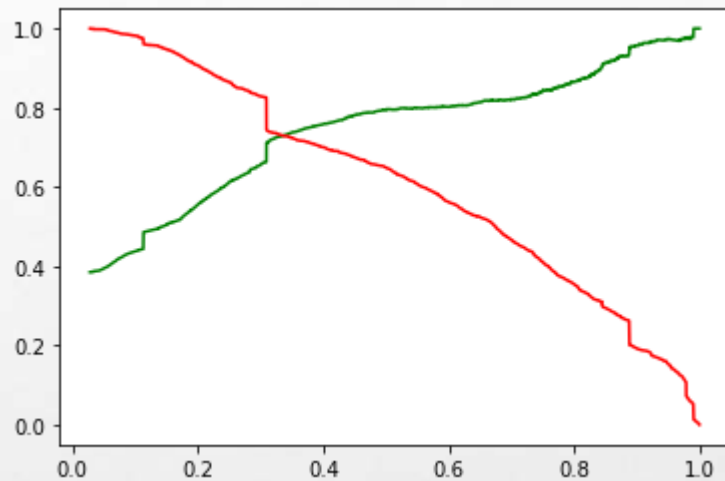
Plotting ROC Curve

The ROC Curve should be a value close to 1. We are getting a good value of 0.86 indicating a good predictive model.



Finding Optimal Cutoff Point

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



Precision and Recall Trade-off

Inference / Conclusion

Model Analysis

Performance of our Final Model

Overall accuracy on Test set: 0.77

Sensitivity of our logistic regression model: 0.83

Specificity of our logistic regression model: 0.74

Conclusion (Recommendation)

Important features responsible for good conversion rate or the ones' which contributes more 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

Thank-you