

Lead Scoring – Case Study

M Kasi Viswanadha Rao Gupta

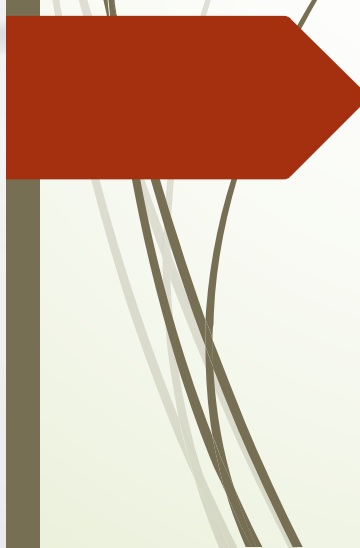
Ramamoorthy Kalimuthu

Mousumi Shaw



Problem Statement

X Education promotes its courses on various websites and search engines, including Google, to attract potential customers. Visitors to the website have the option to browse courses, complete forms, or watch educational videos. Leads are identified when individuals fill out forms, providing their contact details such as email addresses or phone numbers. Additionally, the company receives leads through referrals. Although the sales team engages with these leads through calls, emails, and other means, only a portion of them ultimately convert into paying customers. Currently, the typical lead conversion rate at X Education is approximately 30%.



X Education generating a considerable number of leads, the lead conversion rate remains disappointingly low. For instance, if the company acquires 100 leads in a day, only about 30 of them actually convert into paying customers. To enhance efficiency, X Education aims to identify the most promising leads, commonly referred to as 'Hot Leads.' By successfully identifying and prioritizing these leads, the company expects the lead conversion rate to improve. With a more targeted approach, the sales team can focus their efforts on communicating and nurturing potential leads rather than engaging with every lead indiscriminately. This optimization is crucial to achieving the CEO's target lead conversion rate of approximately 80%.

Objective

The objective of this project is to develop a lead scoring model that assigns a score to each lead, enabling X Education to identify the leads with the highest conversion potential. By implementing this model, the company aims to improve the lead conversion rate and maximize the efficiency of the sales team's efforts. The target set by the CEO is to achieve a lead conversion rate of approximately 80%.

Approach

The proposed approach involves leveraging customer data and lead behavior to assign lead scores that reflect the likelihood of conversion. The model will consider various factors, including lead source, engagement with website content, form completion, and past referrals. By prioritizing leads with higher scores, X Education can optimize its sales efforts by focusing on the most promising leads, enhancing the chances of conversion and revenue growth. The subsequent slides will delve into the details of the lead scoring model and the steps involved in its implementation.

Steps used

Data Cleaning,
Manipulation &
Outlier Analysis

EDA

Data Preparation

Model Prediction

Model Evaluation

Lead score
calculation

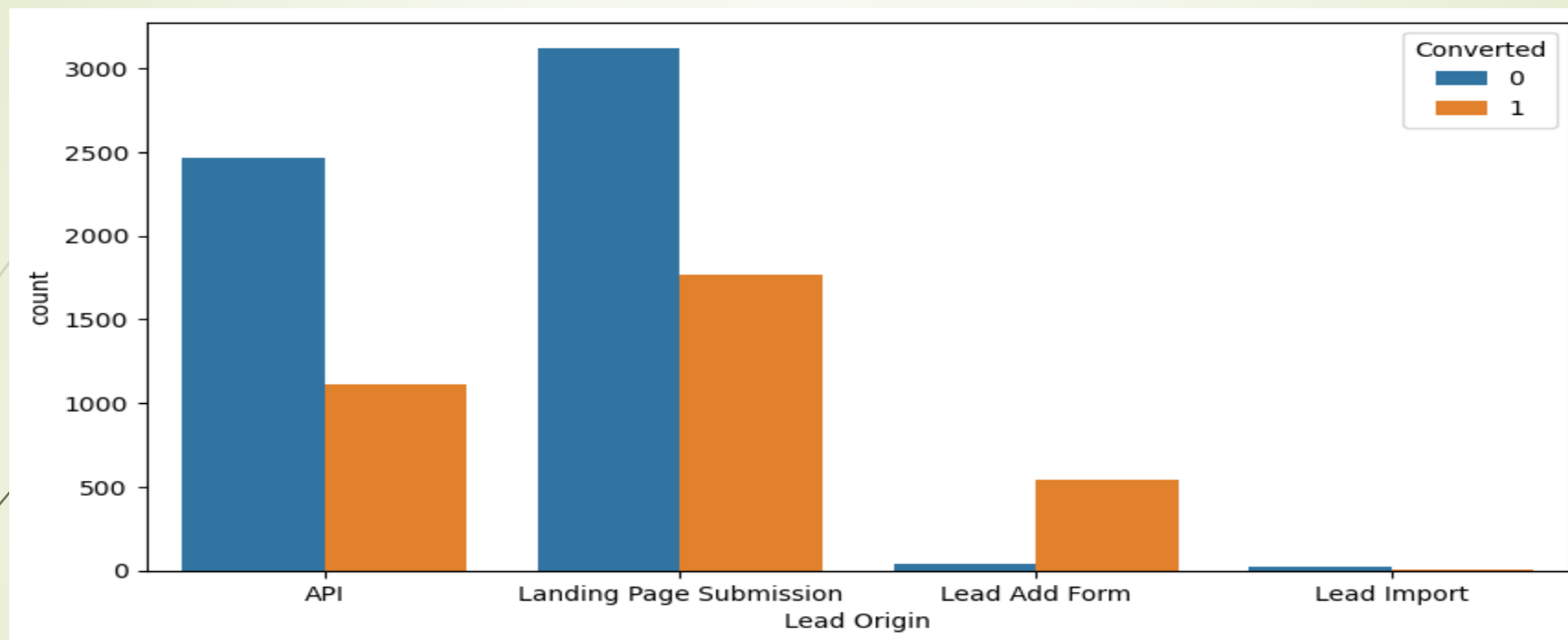
Hot Leads
Determination

Feature
Importance
Determination

Business Insights

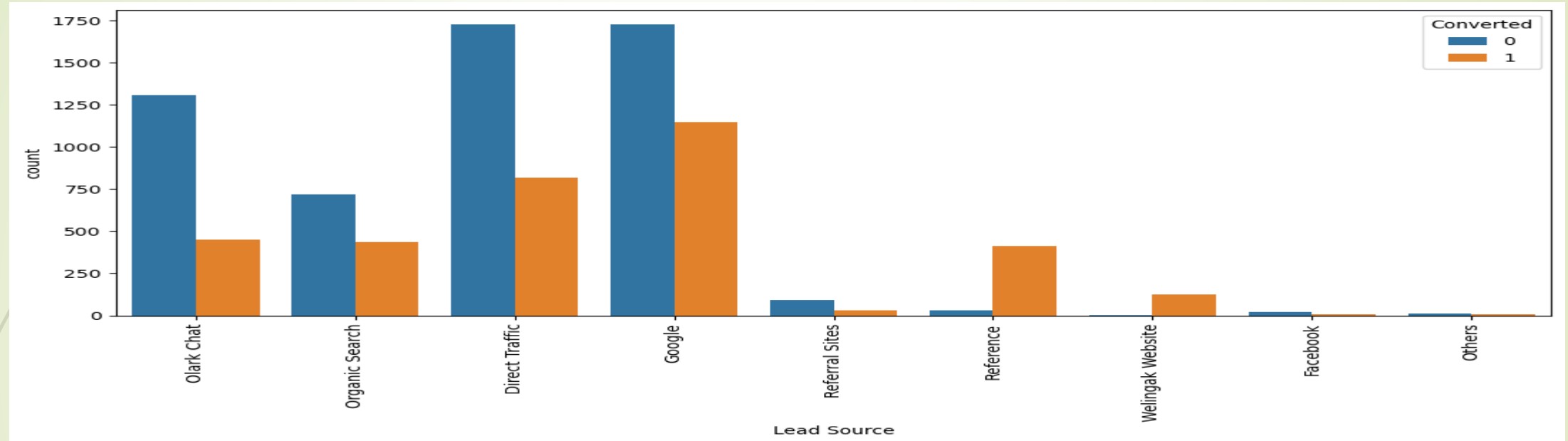
Results &
Recommendation

Exploratory Data Analysis



1. 'API' and 'Landing Page Submission' have 30-35% conversion rate and the count of lead originated from them are considerable.
 2. 'Lead Add Form' has more than 90% conversion rate but count of lead are not very high.
 3. Lead Import are very less in count.
- So, To improve overall lead conversion rate, we need to focus more on improving lead conversion of API and Landing Page Submission origin and generate more leads from Lead Add Form.

Lead Source

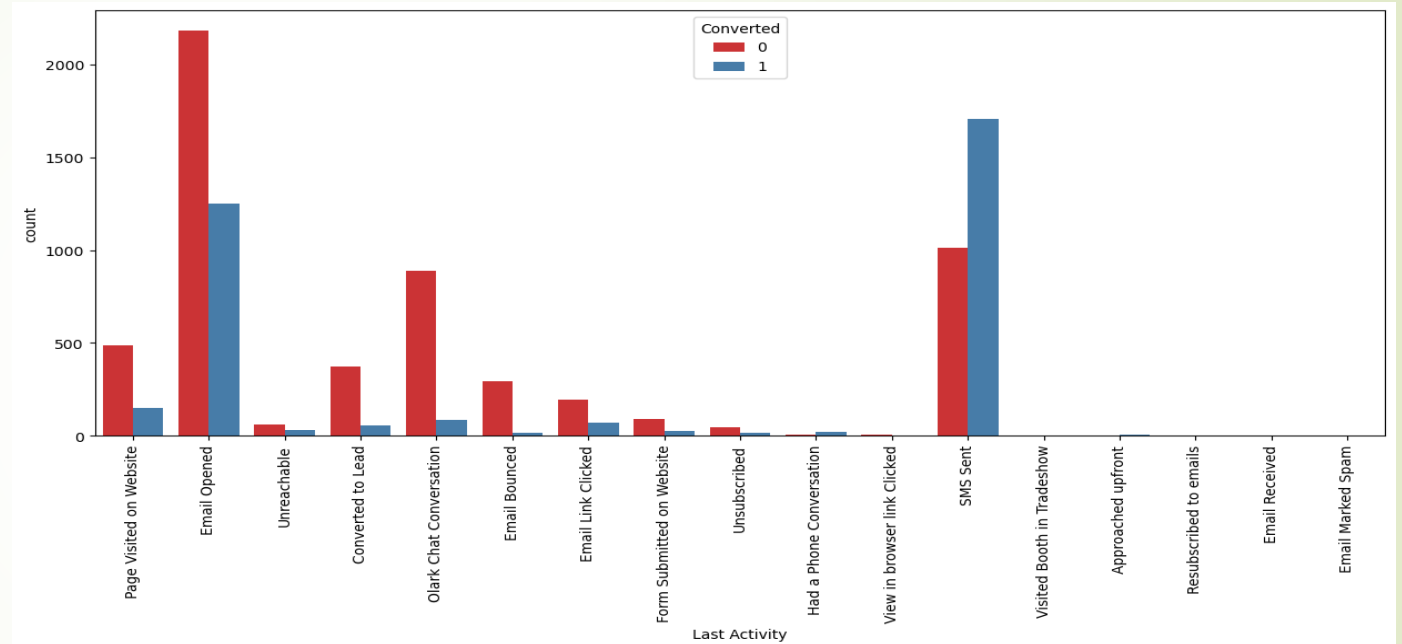


1. Google and Direct traffic generates maximum number of leads.
2. Conversion Rate of reference leads and leads through welingak website is high.

To improve overall lead conversion rate, focus should be on improving lead conversion of olark chat, organic search, direct traffic, and google leads and generate more leads from reference and welingak website.

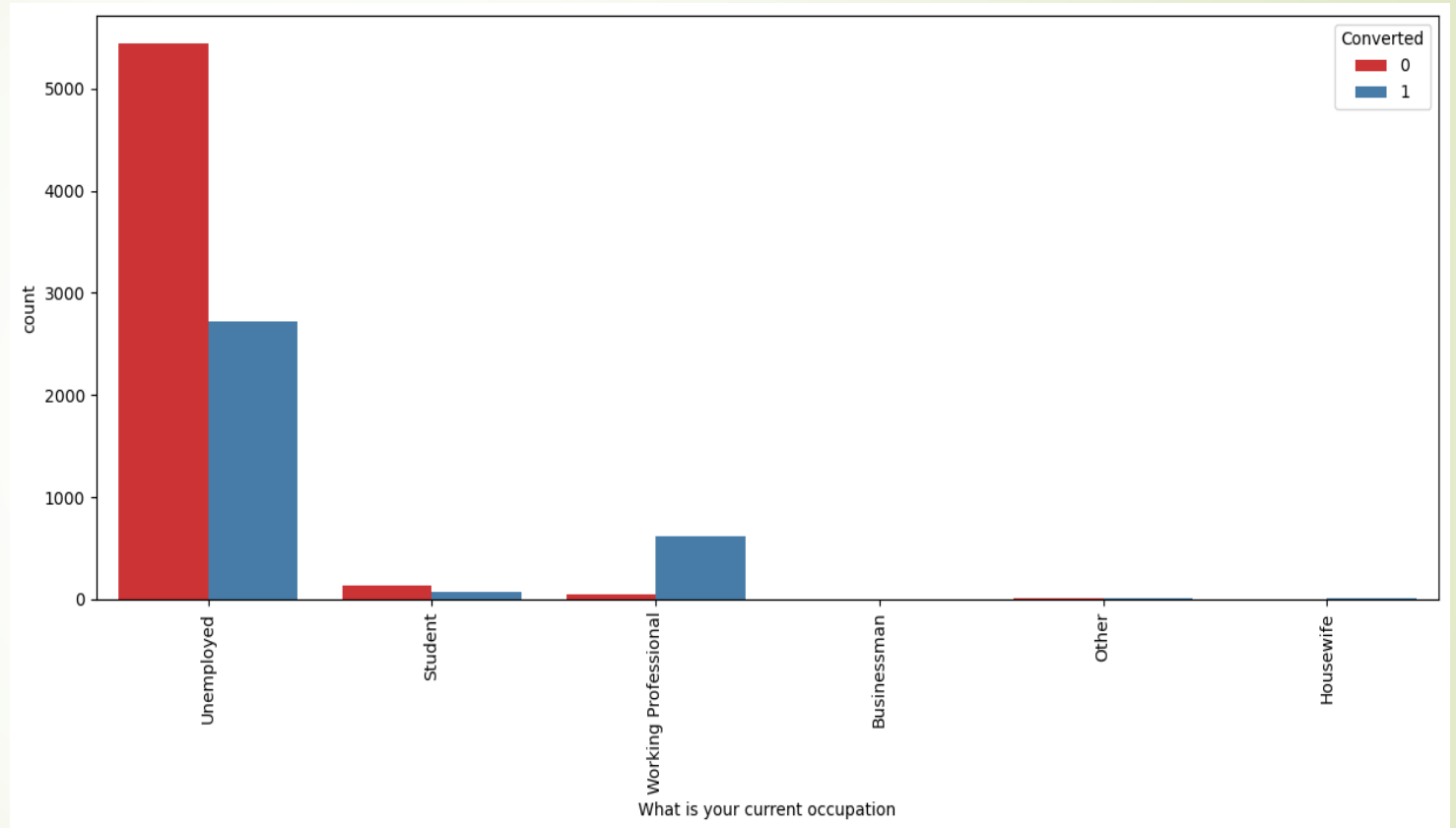
Last Activity

- 1. Most of the lead have their Email opened as their last activity.
- 2. Conversion rate for leads with last activity as SMS Sent is almost 60%.



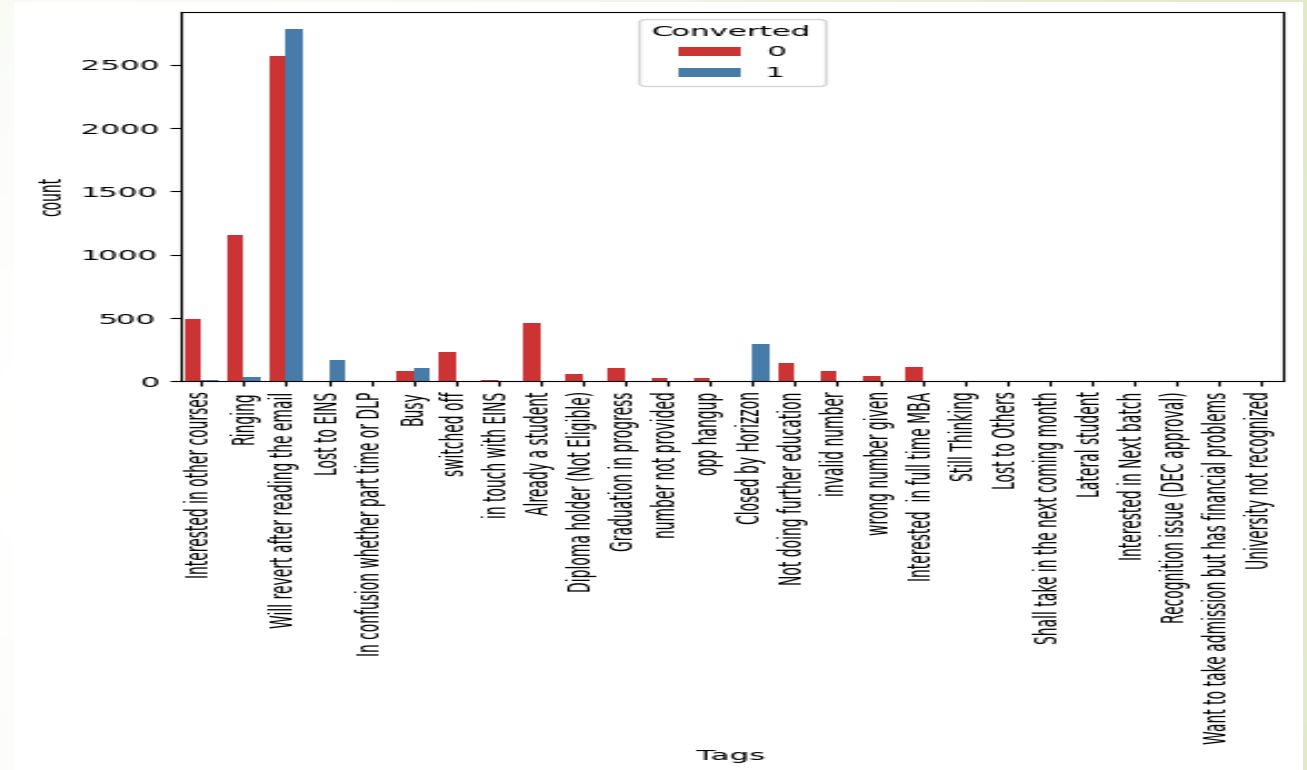
What is your current occupation

➤ Working Professionals going for the course have high chances of joining it. Unemployed leads are the most in numbers but has around 30-35% conversion rate.



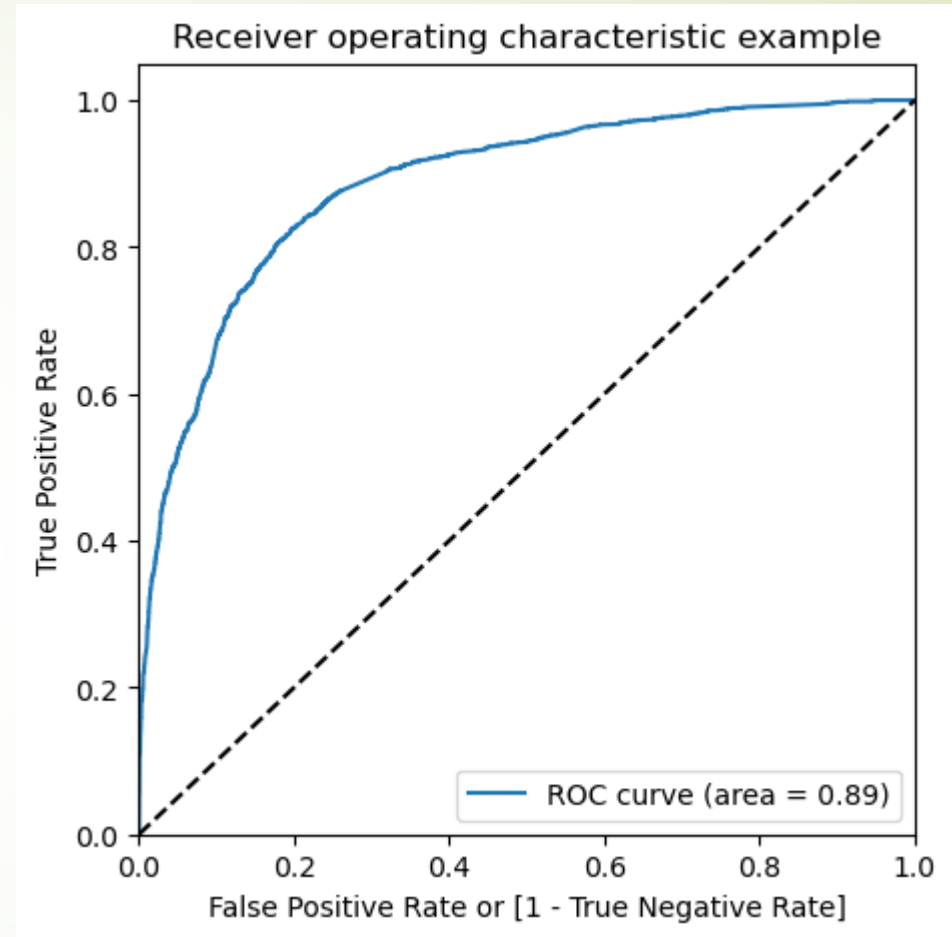
Tags

➤ Working Professionals going for the course have high chances of joining it. Unemployed leads are the most in numbers but has around 30-35% conversion rate.

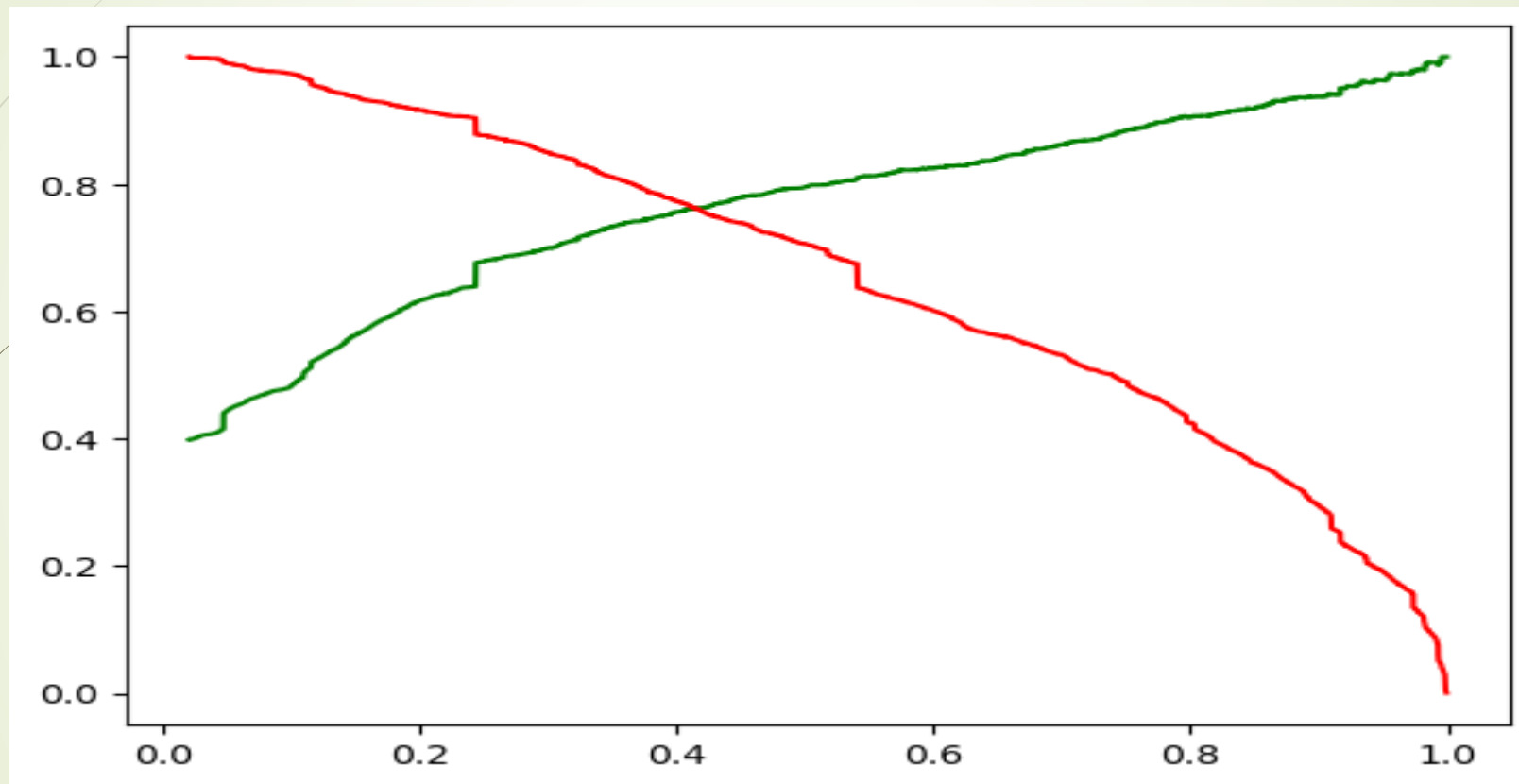


Plotting the ROC Curve

Since we have higher (0.89) area under the ROC curve, therefore our model is a good one.

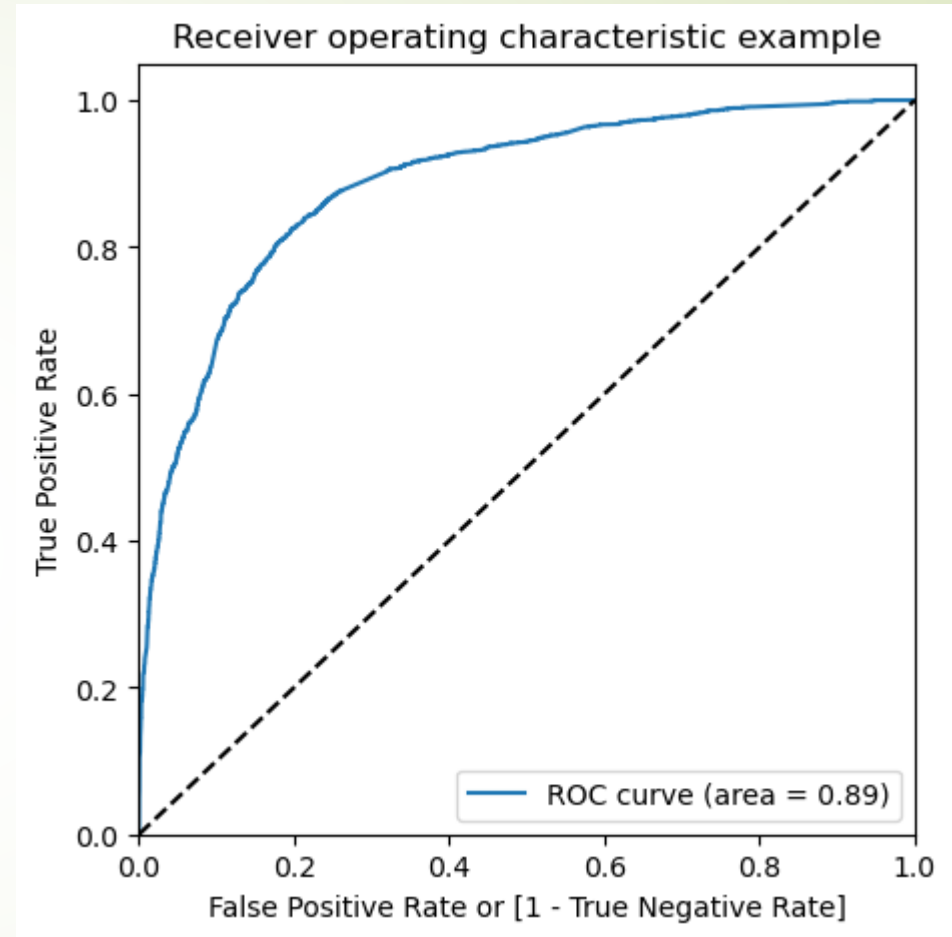


Trade-off curve between precision and recall



Plotting the ROC Curve

Since we have higher (0.89) area under the ROC curve, therefore our model is a good one.



Model Evaluation

- ❖ Confusion matrix was made and cut off point of 0.345 was selected based on accuracy, sensitivity and specificity plot. This cut off gave accuracy, specificity and precision all around 80%. Whereas precision recall view gave less performance metrics around 75%.
- ❖ As to solve business problem CEO asked to boost conversion rate to 80%, but metrics dropped when we took precision-recall view. So, we will choose sensitivity-specificity view for our optimal cut-off for final predictions
- ❖ Lead score was assigned to train data using 0.345 as cut off.

Observations:

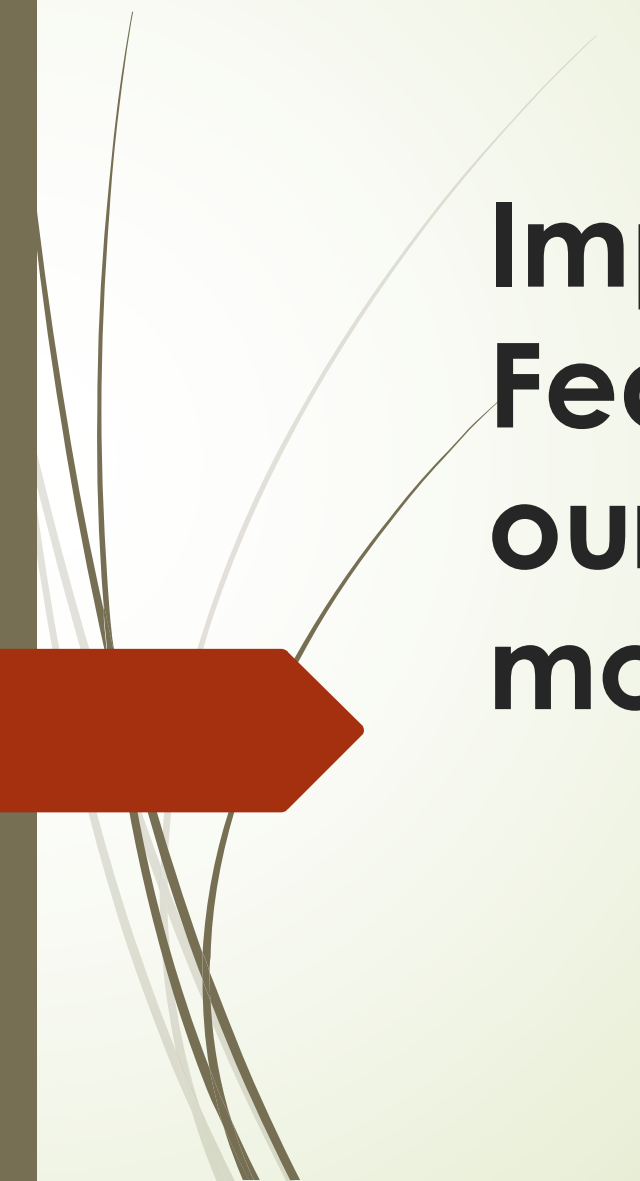
After running the model on the Test Data , we obtain:

Accuracy : 80.4 %

Sensitivity : 80.4 %

Specificity : 80.5 %

Lead Source_Welingak Website	5.811465
Lead Source_Reference	3.316598
What is your current occupation_Working Professional	2.608292
Last Activity_Other_Activity	2.175096
Last Activity_SMS Sent	1.29418
Total Time Spent on Website	1.095412
Lead Source_Olark Chat	1.081908
const	-0.037565
Last Notable Activity_Modified	-0.900449
Last Activity_Olark Chat Conversation	-0.961276
Lead Origin_Landing Page Submission	-1.193957
Specialization_Others	-1.202474
Do Not Email	-1.521825
dtype	float64



Important Features of our final model

Results

- ❖ Comparing the values obtained for Train & Test: Train Data: Accuracy : 81.0 % Sensitivity : 81.7 % Specificity : 80.6 % Test Data: Accuracy : 80.4 % Sensitivity : 80.4 % Specificity : 80.5 % Thus we have achieved our goal of getting a ballpark of the target lead conversion rate to be around 80% . The Model seems to predict the Conversion Rate very well and we should be able to give the CEO confidence in making good calls based on this model to get a higher lead conversion rate of 80%.
- ❖ Finding out the leads which should be contacted: The customers which should be contacted are the customers whose "Lead Score" is equal to or greater than 85. They can be termed as 'Hot Leads'.

Recommendations

- ❖ More budget/spend can be done on Welingak Website in terms of advertising, etc.
- ❖ Incentives/discounts for providing reference that convert to lead, encourage to provide more references.
- ❖ Working professionals to be aggressively targeted as they have high conversion rate and will have better financial situation to pay higher fees too.

Thank You !!!

