

X EDUCATION- LEAD SCORING CASE STUDY

From:

Vishwas aithal

Background

- X Education, an education company which sells online courses to industry professionals.
- Many professionals who are interested in such courses go to their websites and then browse for desired courses.
- On several websites and search engines like Google, the company does its marketing
- Once these people land on the website, they might browse the courses or fill up a form for the course or watch some videos

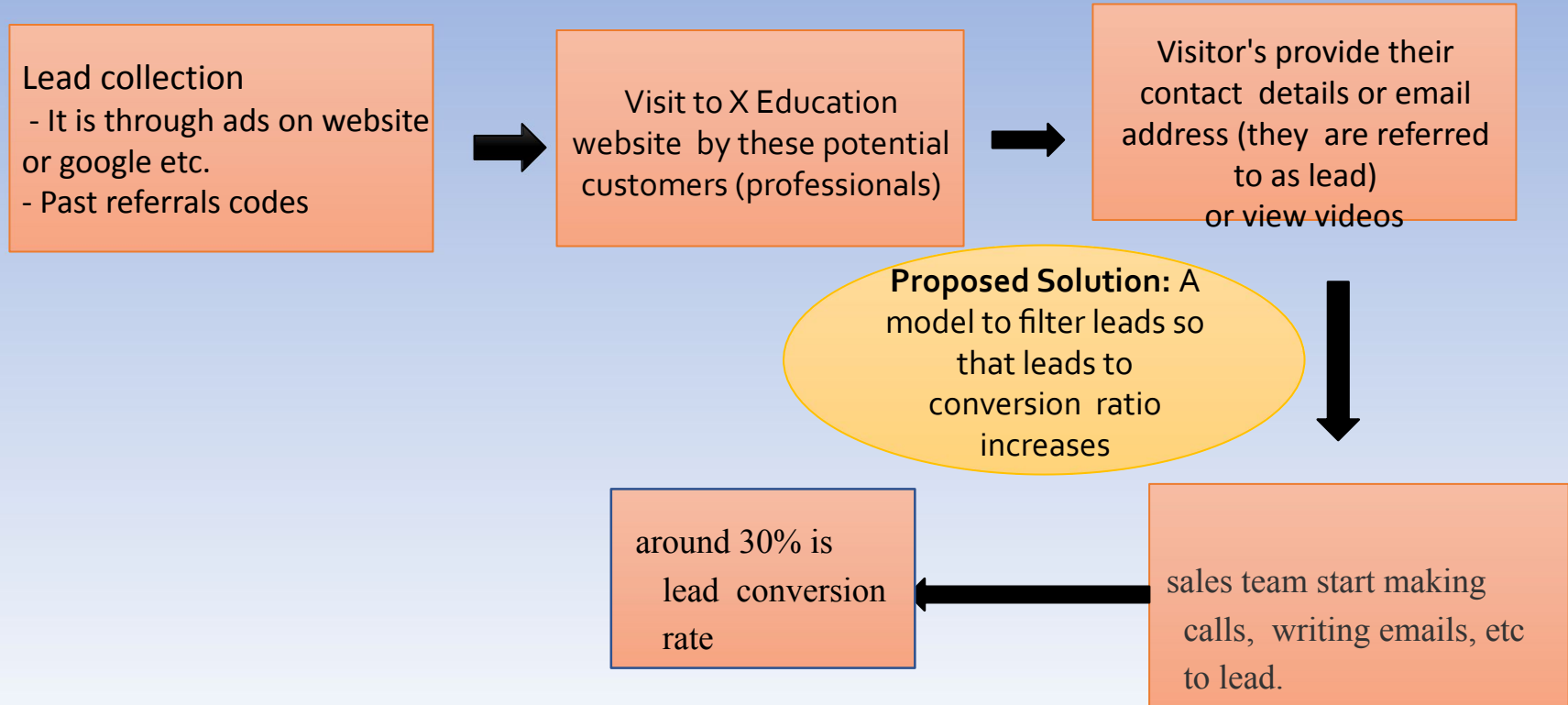
- When these people fill up a form providing their email address or phone number, they are classified to be a lead.
- The company also gets leads through past referrals.
- Once these leads are acquired, employees from the sales team start making calls, writing emails, etc to them.
- Through this process, some of the leads get converted while most do not. The typical lead conversion rate at X education is around 30%.

Problem Statement

- Although X Education gets a lot of leads, 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 it is successfully identify this set of leads or the lead conversion rate should go up.
- And the sales team will be focusing more on communicating with the potential leads rather than making calls to everyone.

- We need to help them to select the most promising leads, that is by selecting the leads that are most likely to convert into paying customers.
- We need 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 Rate



Proposed Solution

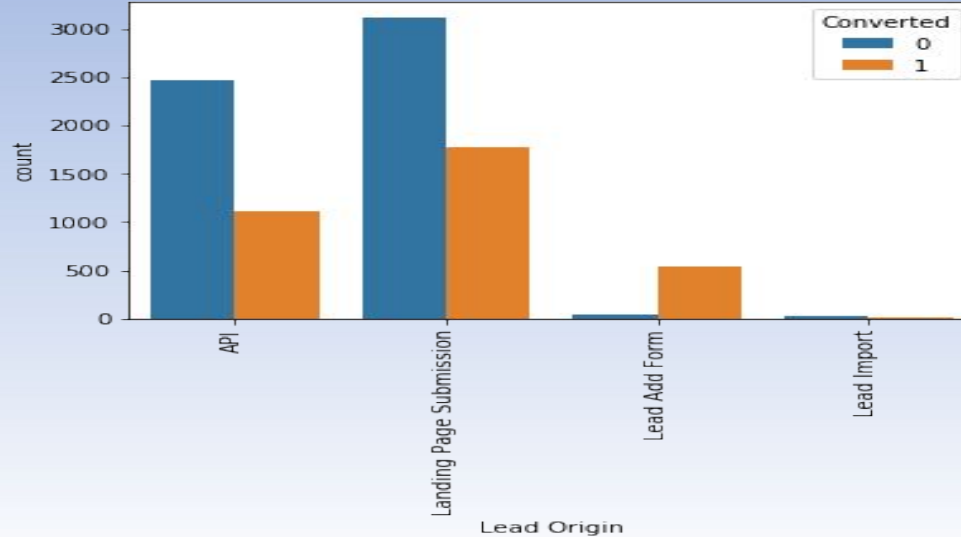
- **Selection of Hot leads:** Here we cluster the leads into different categories based on chances to get converted from leads to hot leads so that we can focus more on them.
- **Communication with Hot leads:** Now we have a smaller group of leads sales team can contact them and can convince them .
- **Conversion of Hot Leads:** since there is smaller group of leads we have more focussed and more probable of converting them to hot leads and then we can achieve 80% target.

Steps Involved:

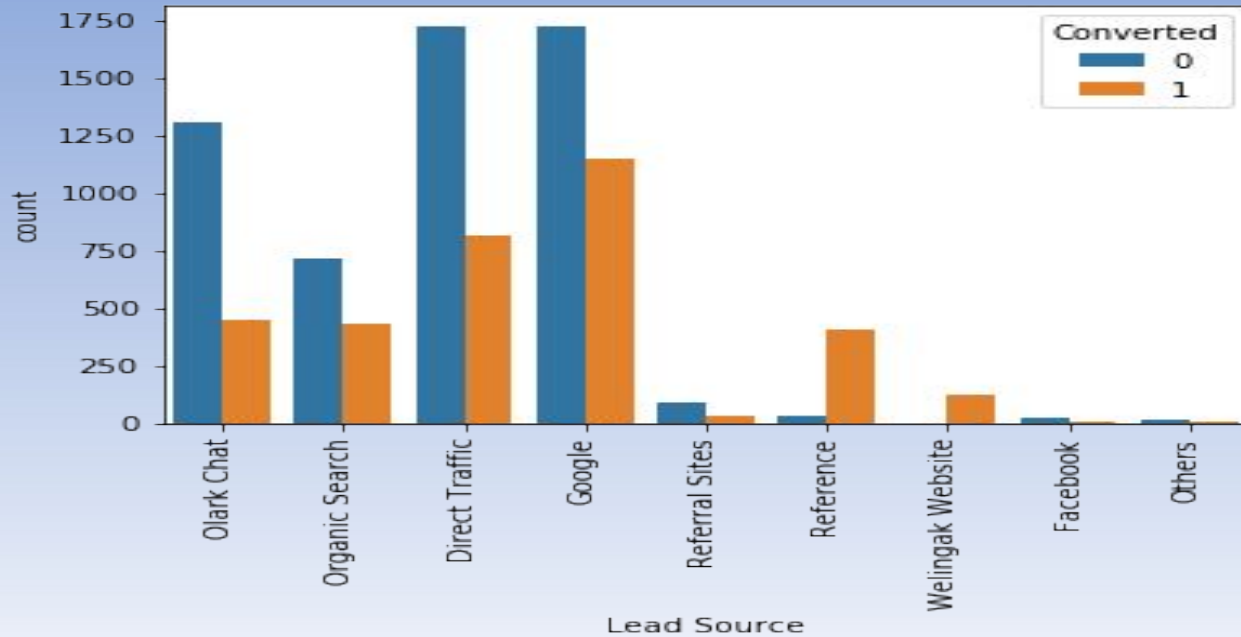
- **Data Collection:** The data which is provided by the company is been loaded, the columns and rows were gone through so that we get clear understating what each column specifies.
- **Data Cleaning :** Here we removed data with duplicate values, high null values, some unnecessary column.
- **Performing EDA:** Different types of analysis is been done like Univariate, Bivariate, and Heatmap for numerical and categorical columns

- Data preparation: The outlier which has higher value or may disturb our analysis was treated either by removing or standardisation.
- Model Building: Model building using RFE for selected columns
- Model Improvement : Reduction of some more columns and Model re-building.
- Final Model : Final Model Analysis and performance on Test Data

Plots(Visualisation)

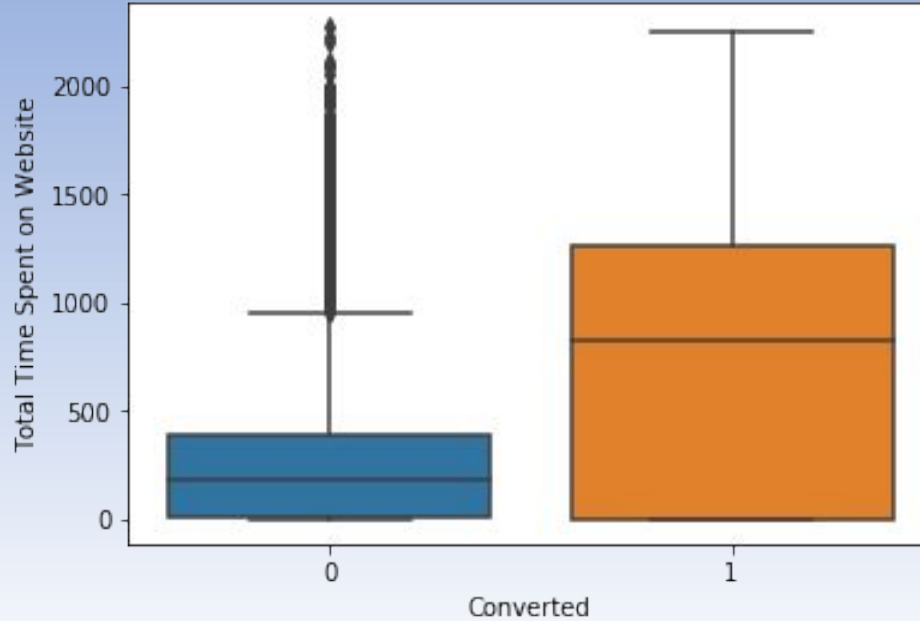


To improve overall lead conversion rate, we need to focus more on improving lead conversion of API and Landing Page Submission origin and generate more leads from Lead Add Form.

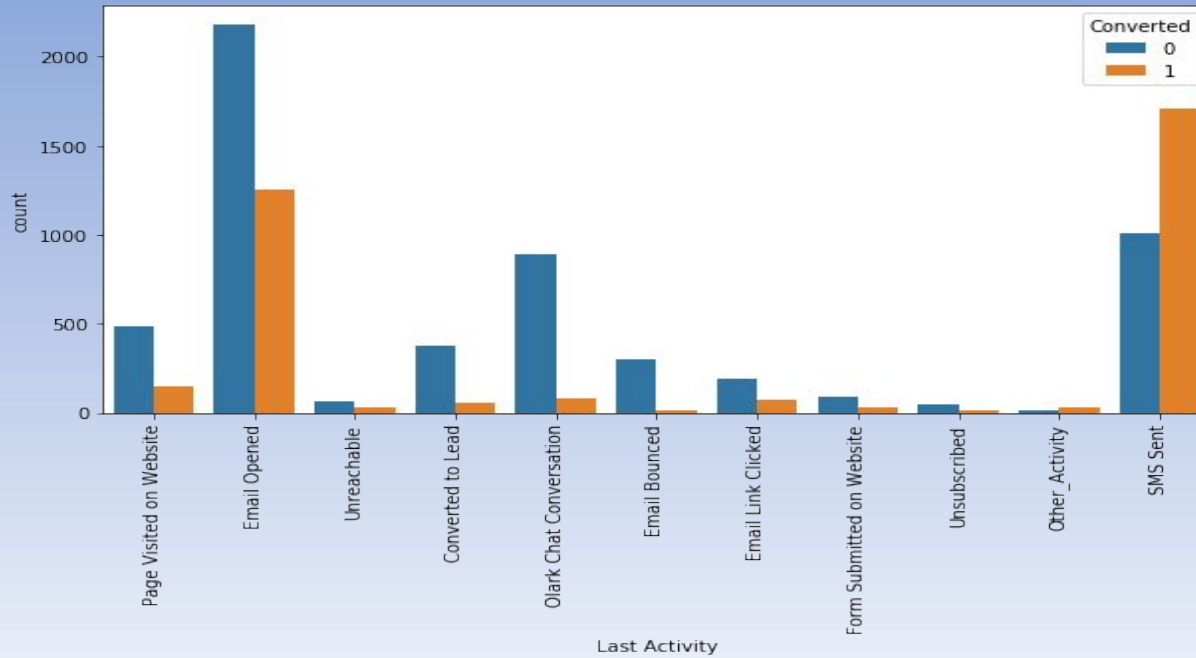


To improve overall lead conversion rate, focus should be on improving lead conversion of olark chat, organic search, direct traffic, and google leads and generate more leads from reference and welingak website.

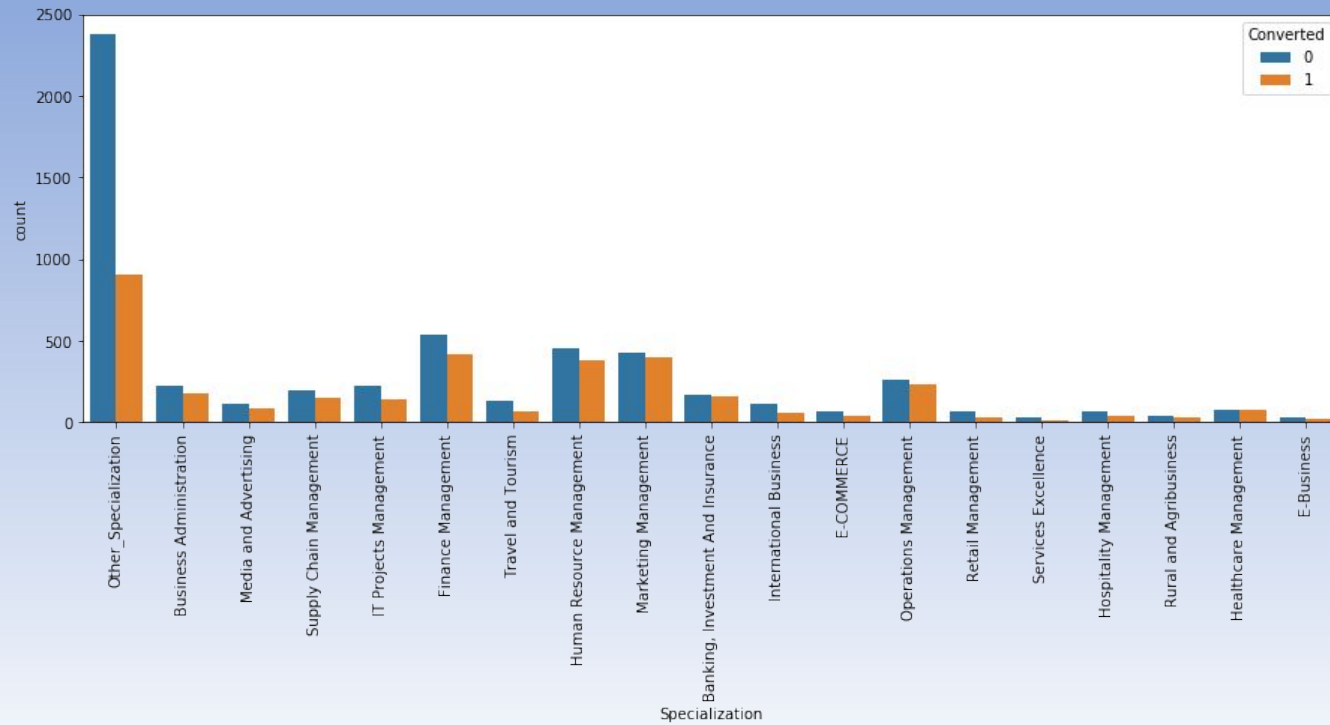
Nothing conclusive can be said on the basis of Total Visits.



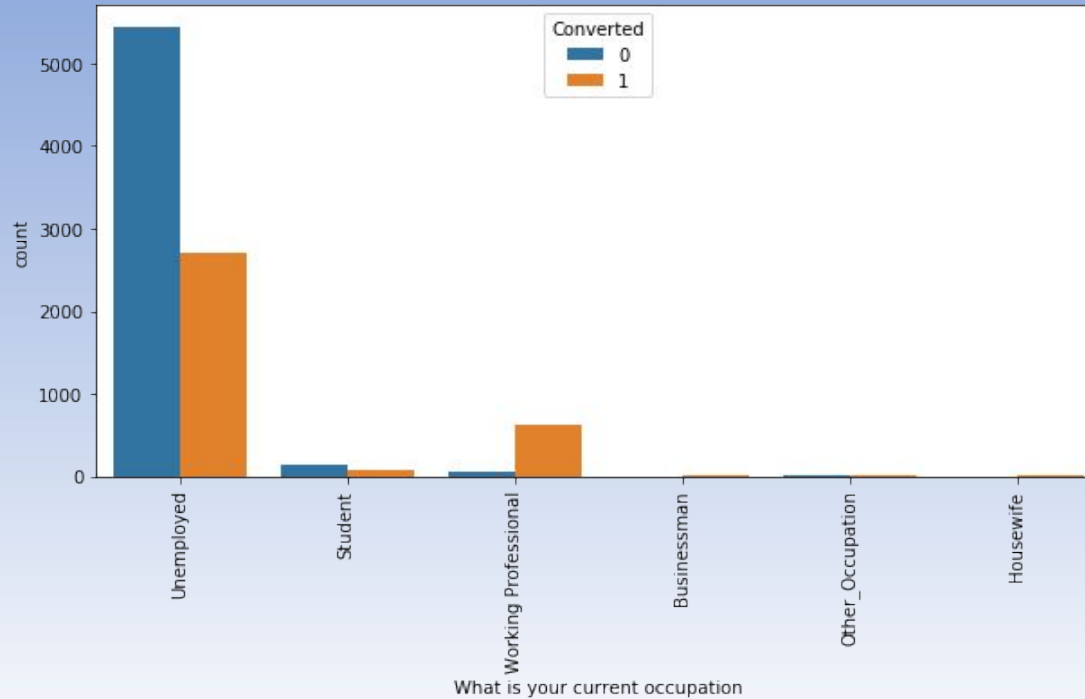
Website should be made more engaging to make leads spend more time.



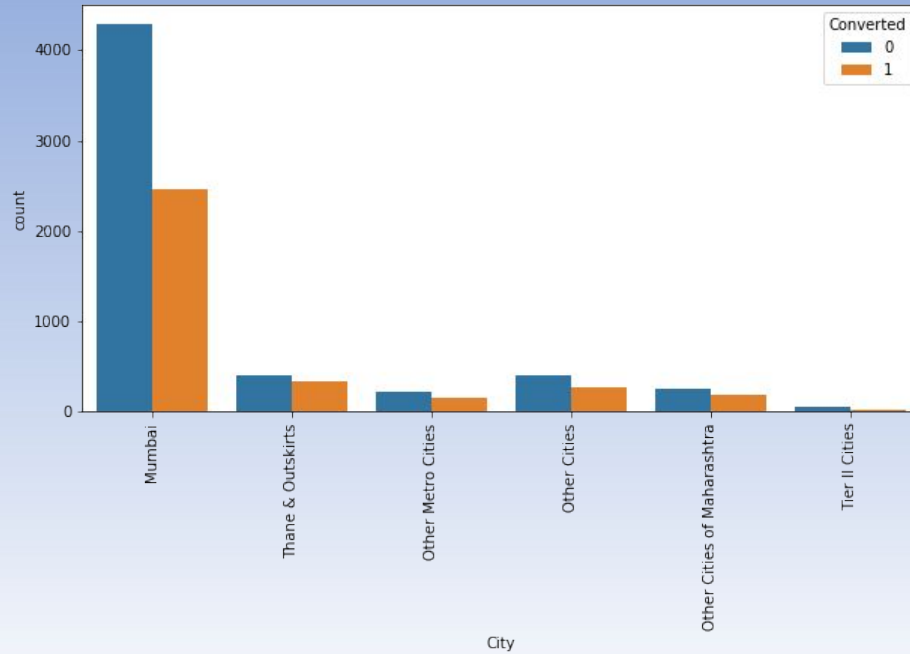
1. Most of the lead have their Email opened as their last activity.
2. Conversion rate for leads with last activity as SMS Sent is almost 60%.



Focus should be more on the Specialization with high conversion rate.

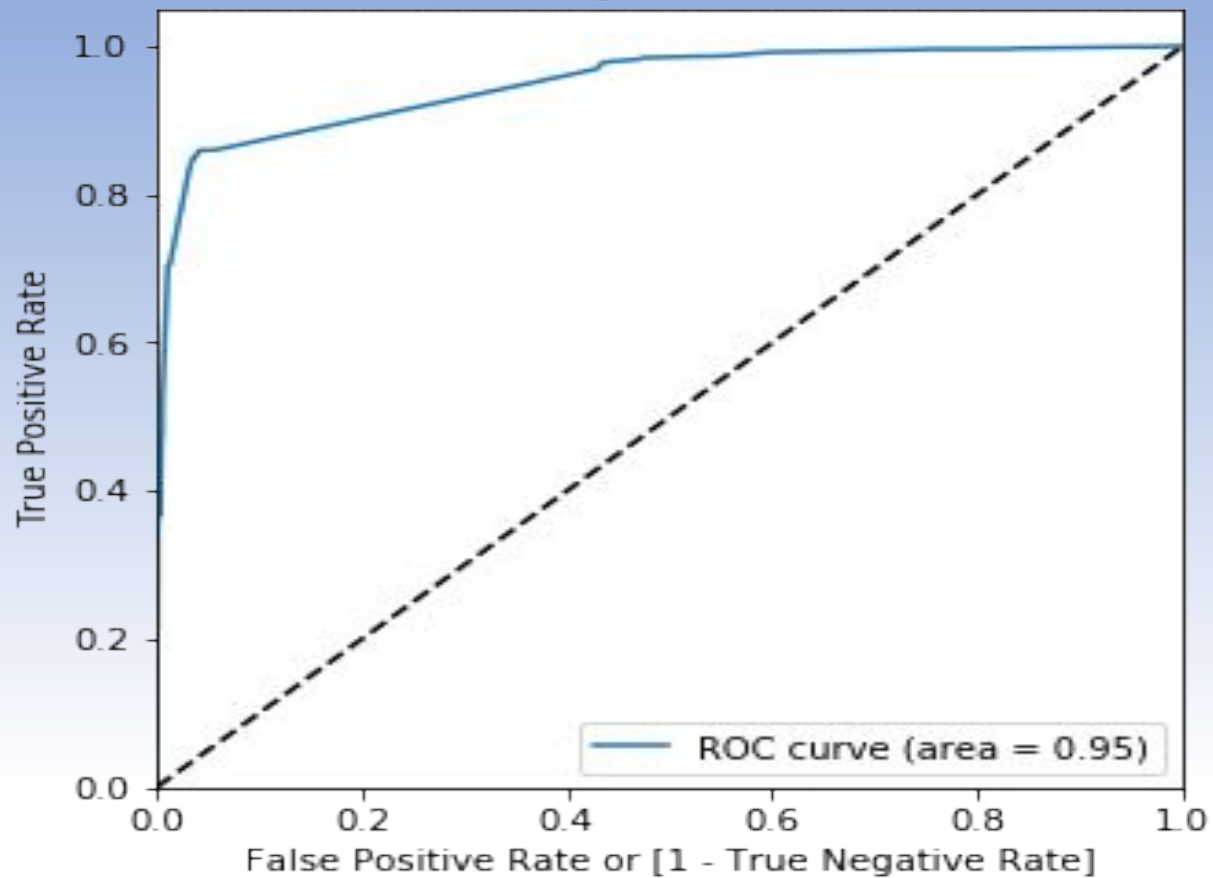


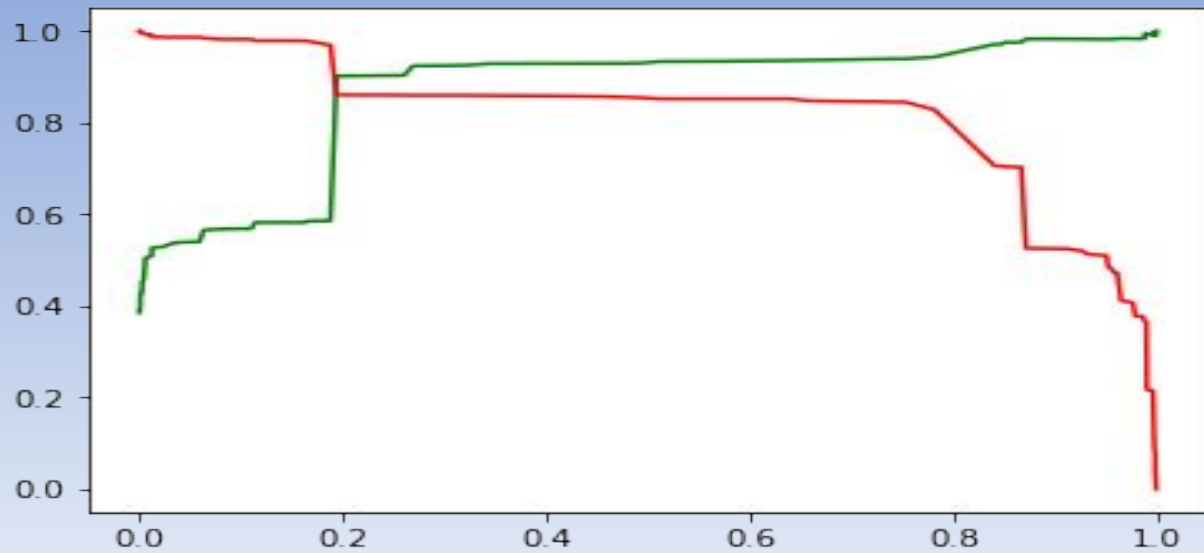
1. Working Professionals going for the course have high chances of joining it.
2. Unemployed leads are the most in numbers but has around 30-35% conversion rate.



Most leads are from mumbai with around 30% conversion rate.

Receiver operating characteristic example





Precision and recall
curve

Model Analysis

Overall accuracy of the model : 90.45

Sensitivity of our logistic regression model:

84.42 Specificity of our logistic regression

model: 93.88

CONCLUSION

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.

*Thank
you*

