

# LEAD SCORING CASE STUDY




Submitted By-  
Anusha Kulkarni  
Husn Ara


# 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.
- The company markets its courses on several websites and search engines 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. Moreover, the company also gets leads through past referrals. 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%.
- although X Education gets a lot of leads, its lead conversion rate is very poor. For example, 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'.





# ANALYSIS APPROACH



- 
- Data Preparation
  - Data Inspection
  - Data Cleaning

- 
- EDA
  - Uni-variate Analysis
  - Data Preparation for Logistic Regression

- 
- Performing Train-Test Split On Data
  - Feature Scaling
  - Model Building

- 
- [text]
- Feature Selection Using RFE
  - Checking VIFs
  - Metrics beyond simply accuracy

- 
- [text]
- Plotting the ROC Curve
  - Finding Optimal Cutoff Point
  - Assigning Lead Score

- 
- [text]
- Precision and recall tradeoff
  - Making predictions on the test set
  - Checking the overall accuracy
- 

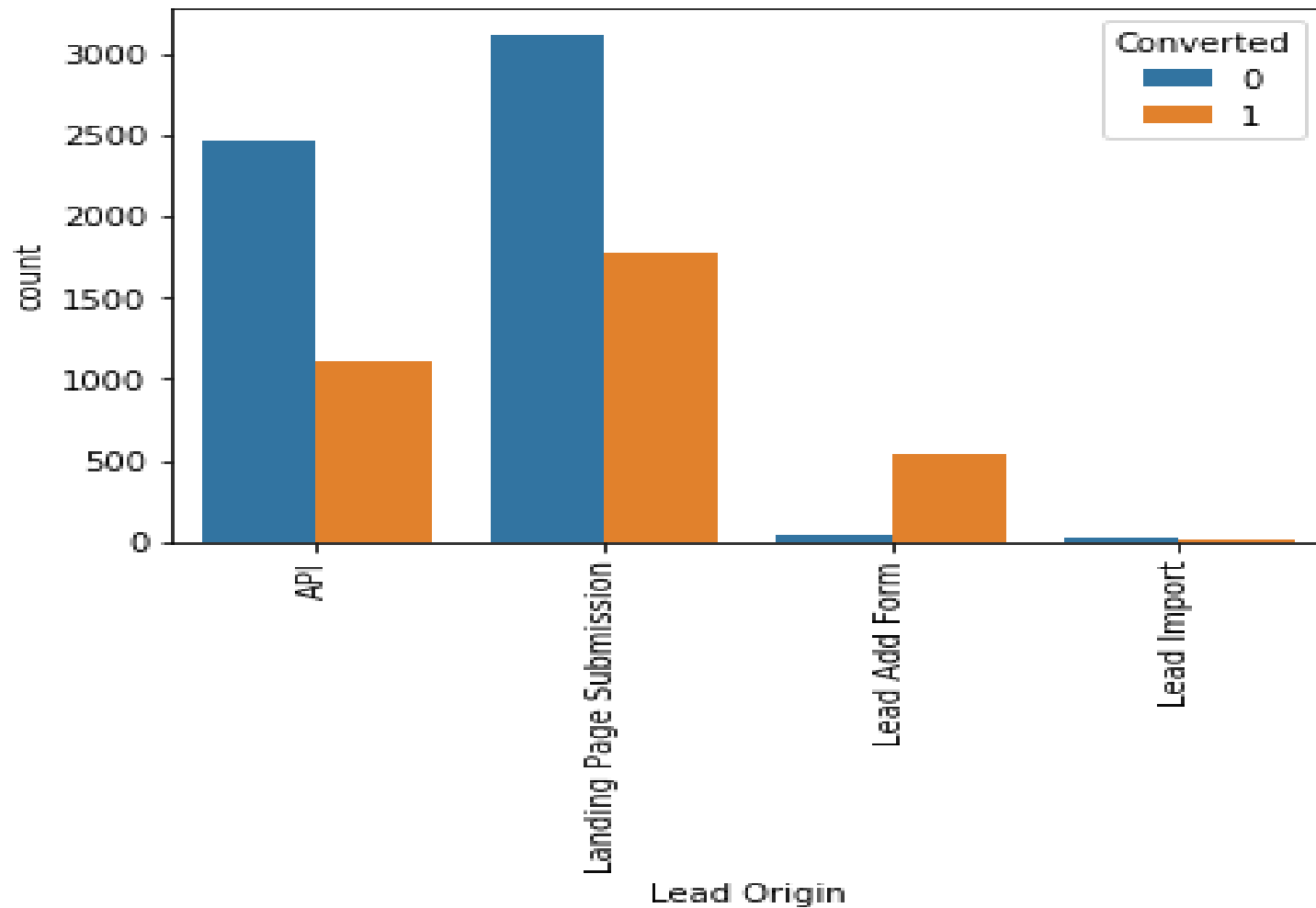
## RESULTS IN BUSINESS TERMS

- We got rate of conversion as 37.85%.
- Most of the leads are converted from Lead originated by **API** and **Landing Page submission**. Very less amount of leads are originated from **lead add form**.
- **Direct Traffic** and **Google** are the major lead sources.
- To improve overall lead conversion rate, focus should be on improving lead conversion of Olark chat, organic search.
- Direct traffic, and Google leads generate more leads from reference and welingak website.
- The rate of conversion of the customers whose last activity is Email\_opened or SMS\_Sent is higher.
- Unemployed leads are more in numbers. Working professionals have high conversions rate.



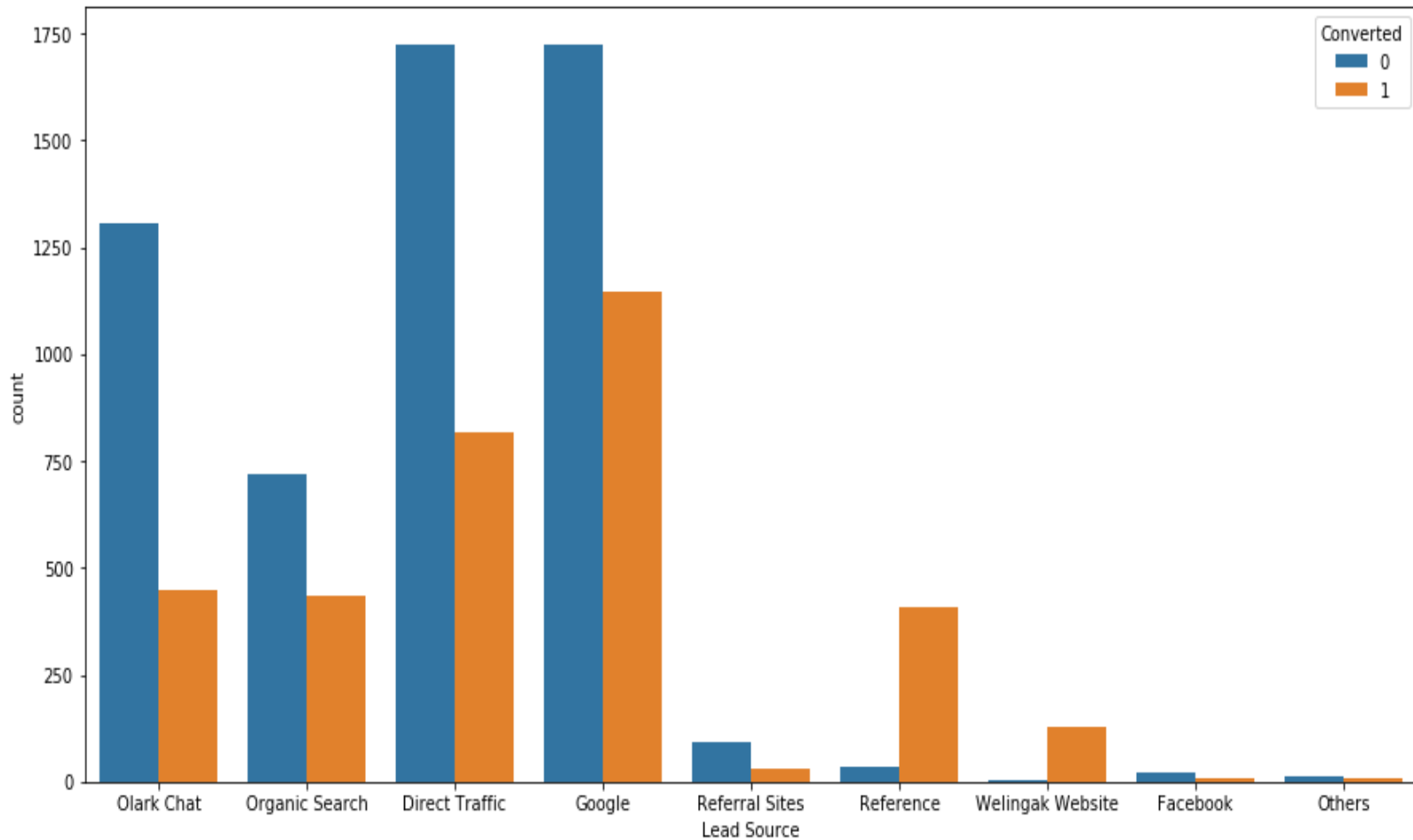
# LEAD ORIGINS

Landing Page Submission and API are higher Lead Origins

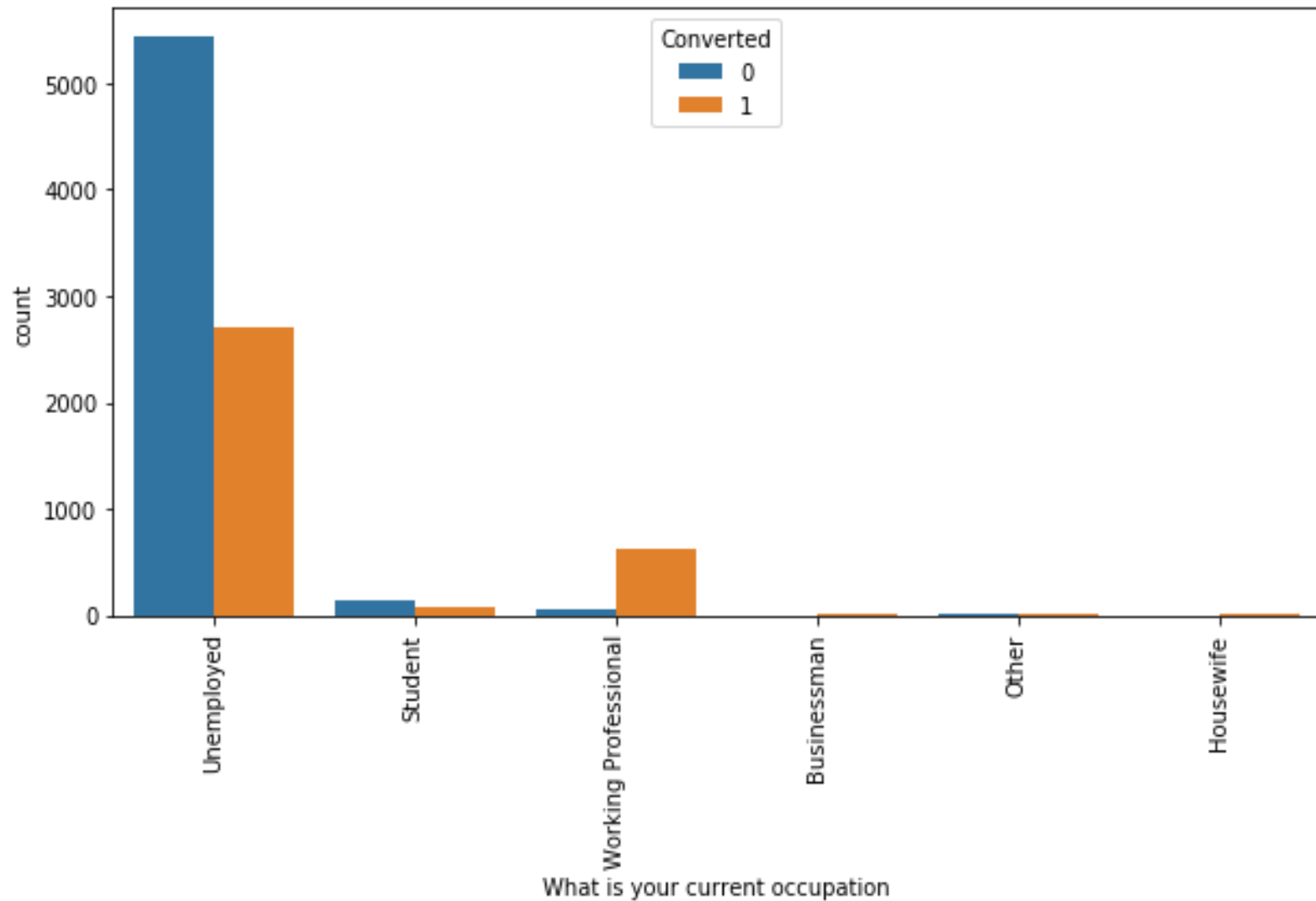


# LEAD SOURCES

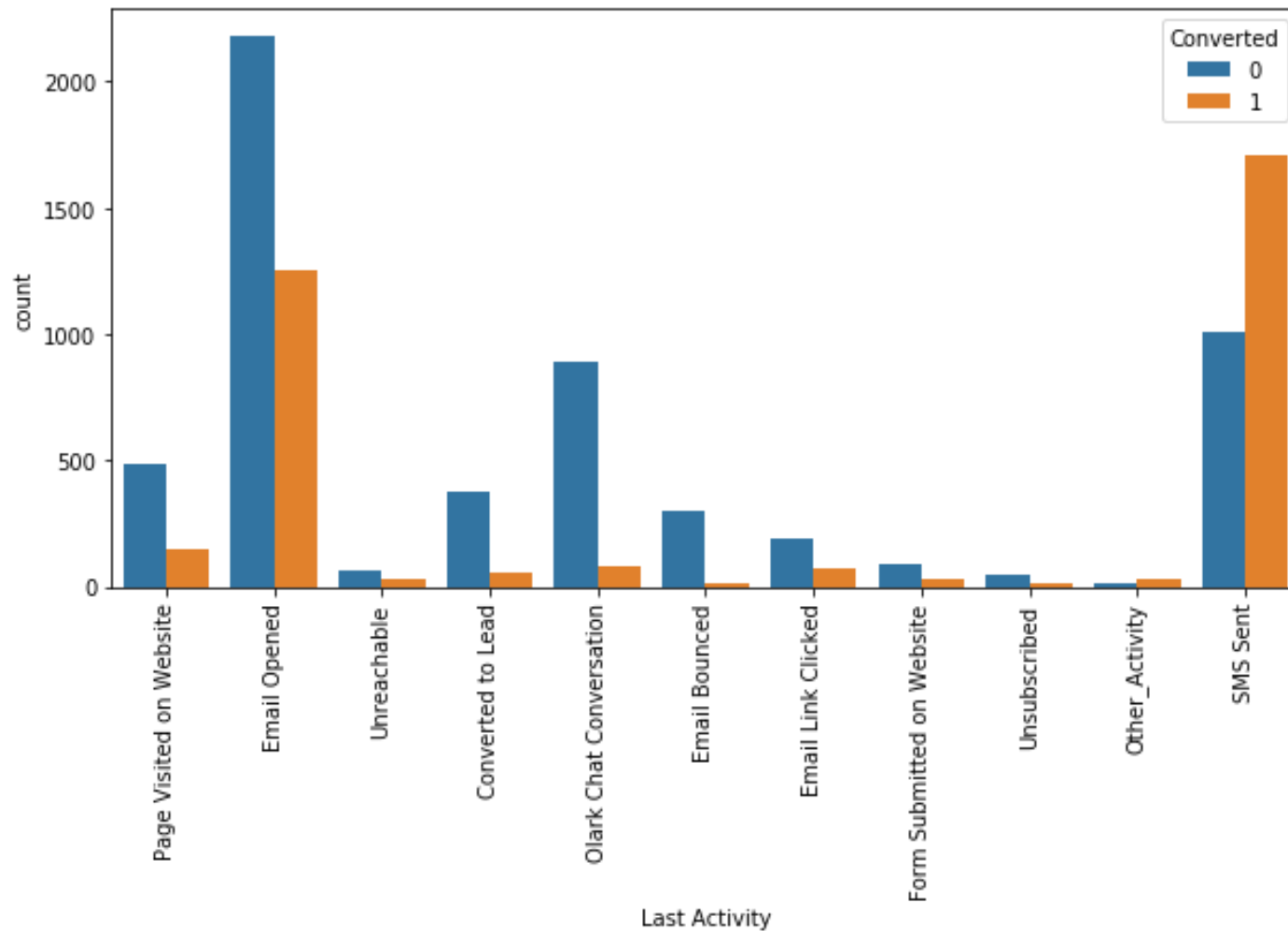
Google and Direct Traffic are higher Lead Sources



UNEMPLOYED LEADS ARE MORE IN NUMBERS.  
WORKING PROFESSIONALS HAVE HIGH CONVERSIONS RATE.

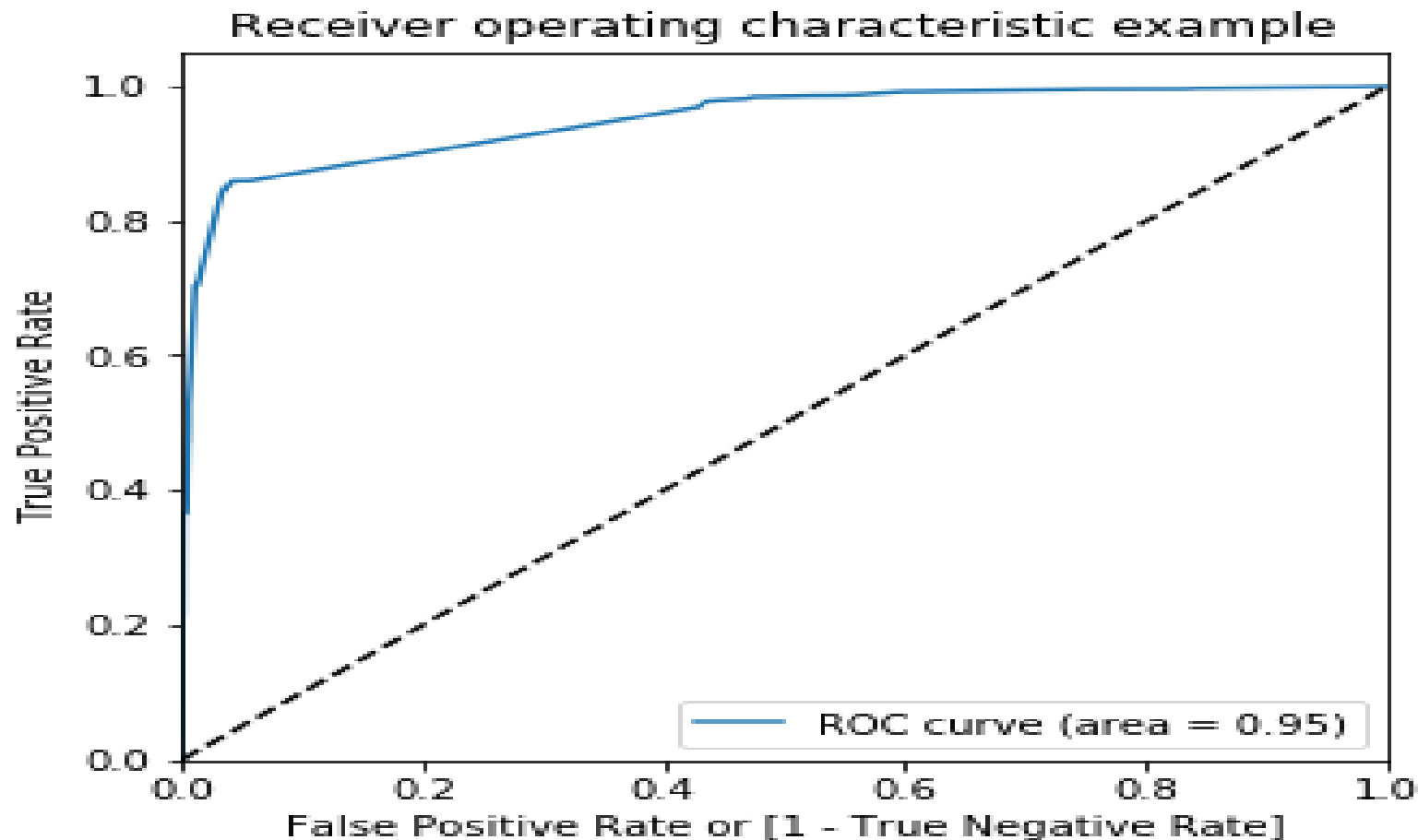


THE RATE OF CONVERSION OF THE CUSTOMERS WHOSE LAST ACTIVITY IS EMAIL\_OPENED OR SMS\_SENT IS HIGHER.





# ROC CURVE



# SUMMARY

- To improve overall lead conversion rate, focus should be on improving lead conversion of Olark chat, Organic search, Direct traffic & Google leads and generate more leads from reference and Welingak website.
- The rate of conversion of the customers whose last activity is Email\_opened or SMS\_Sent is higher. Company must focus on it.
- 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.
- Following are the 3 variables in the model which should be focused the most in order to increase probability of the lead conversion;
  - Lead Origin - Lead add form
  - Lead Source- Welingak Website
  - What is your current occupation - Working Professional



THANK YOU!

