

Lead Scoring Case Study

Submitted by:

Souraneer

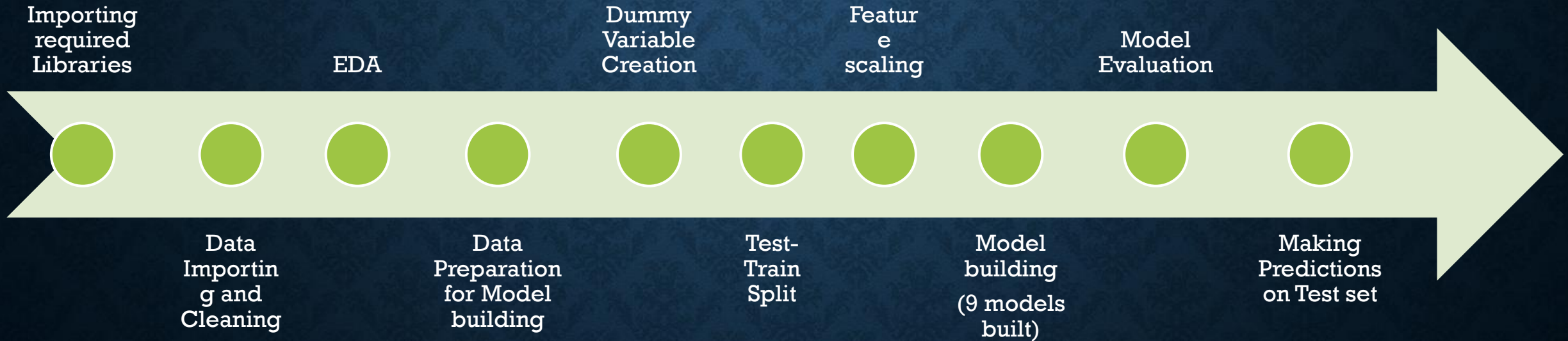
Afsha Shaikh

Sachin

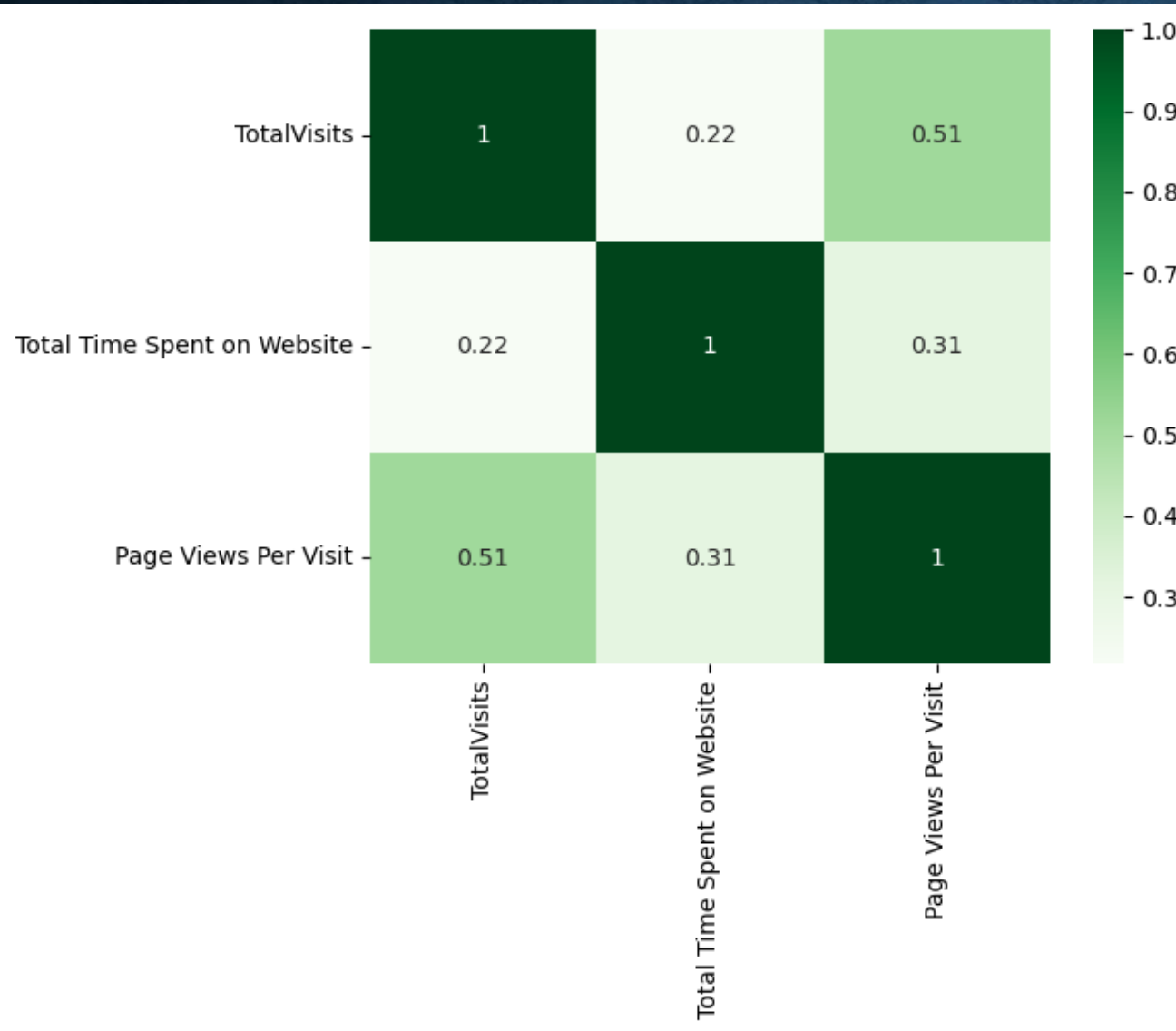
PROBLEM STATEMENT

- An Education company named “X Education”, sells online courses to industry professional.
- The company markets its courses on several websites and search engines. 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.
- Now, although X Education gets a lot of leads, its lead conversion rate is very poor i.e. 30%.
- We are required to build a model wherein every lead is assigned a lead score. Higher the lead score, higher the conversion chance.
- The CEO, in particular, has given a ballpark of the target lead conversion rate to be around 80%.

PROBLEM APPROACH



Numerical features

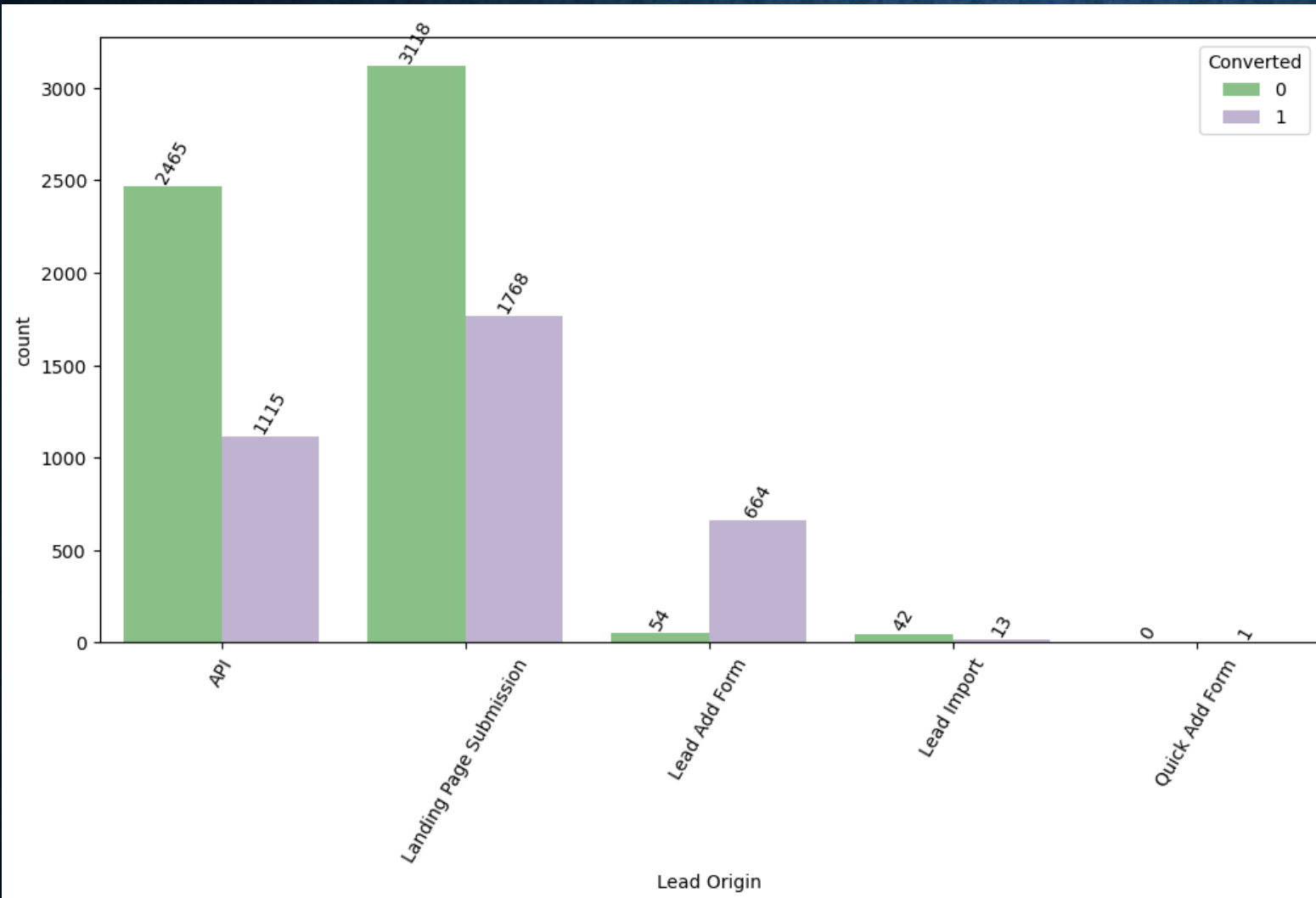


High Correlation between

- Total visits & Page Views per visit

Categorical Features - 1

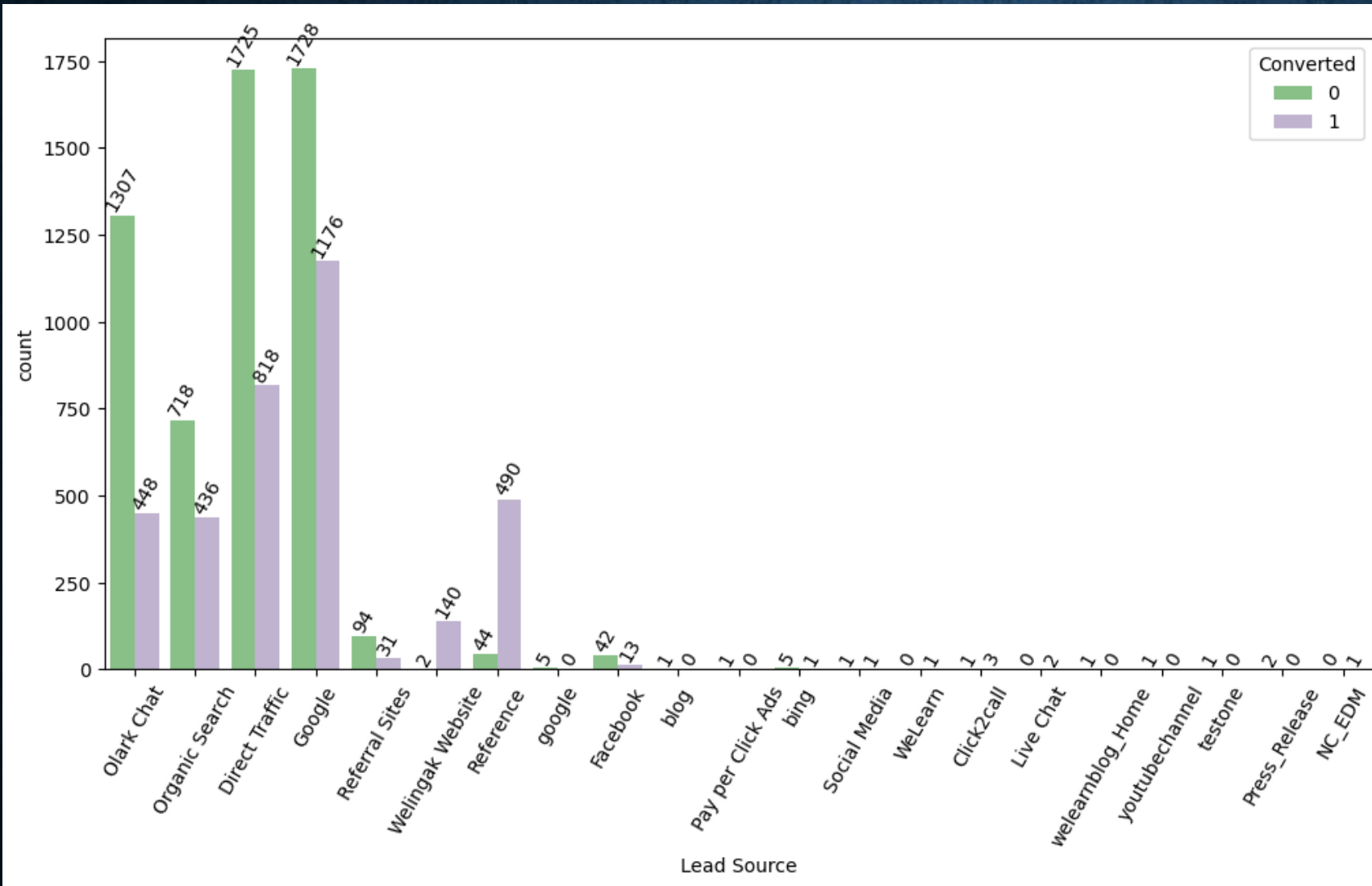
Comparison of leads converted for 'Lead Origin'



'Lead Add Form' has significantly Higher Lead Conversion Rate

Categorical Features - 2

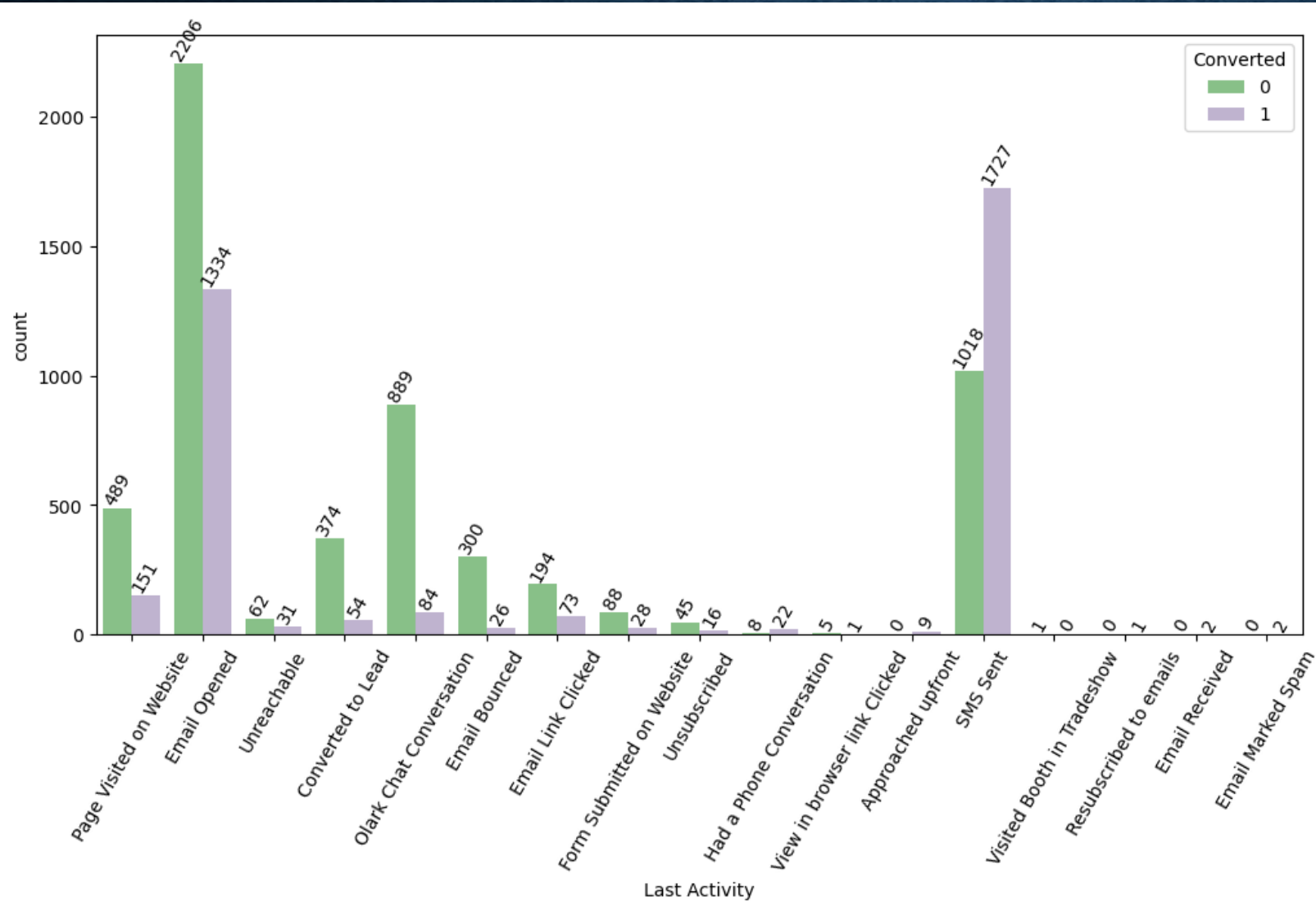
Comparison of leads converted for 'Lead Source'



'Reference', 'Welingak Website' has significantly Higher Lead Conversion Rate

Categorical Features - 3

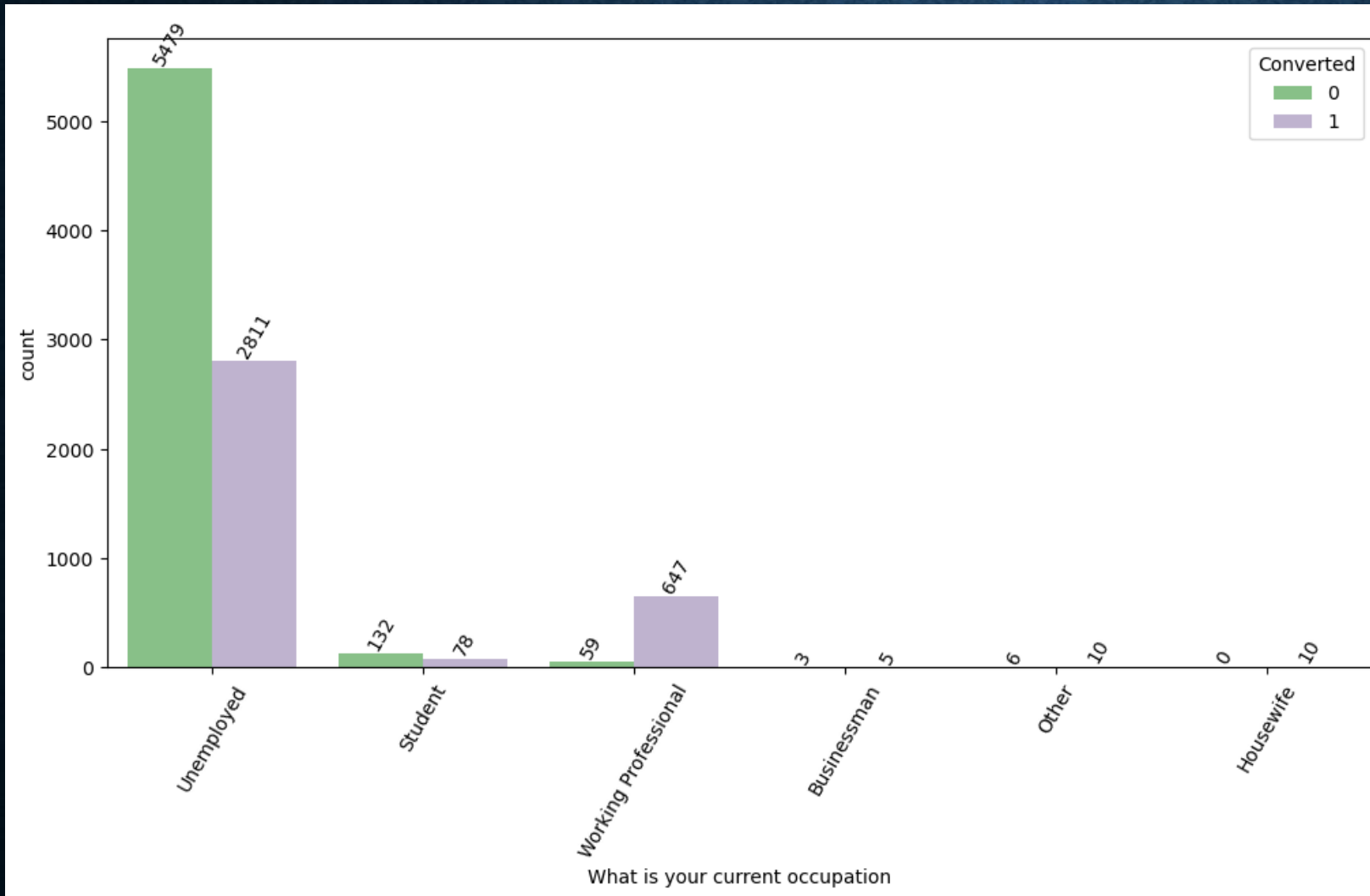
Comparison of leads converted for 'Last Activity'



'SMS sent' has significantly
Higher Lead Conversion Rate

Categorical Features - 4

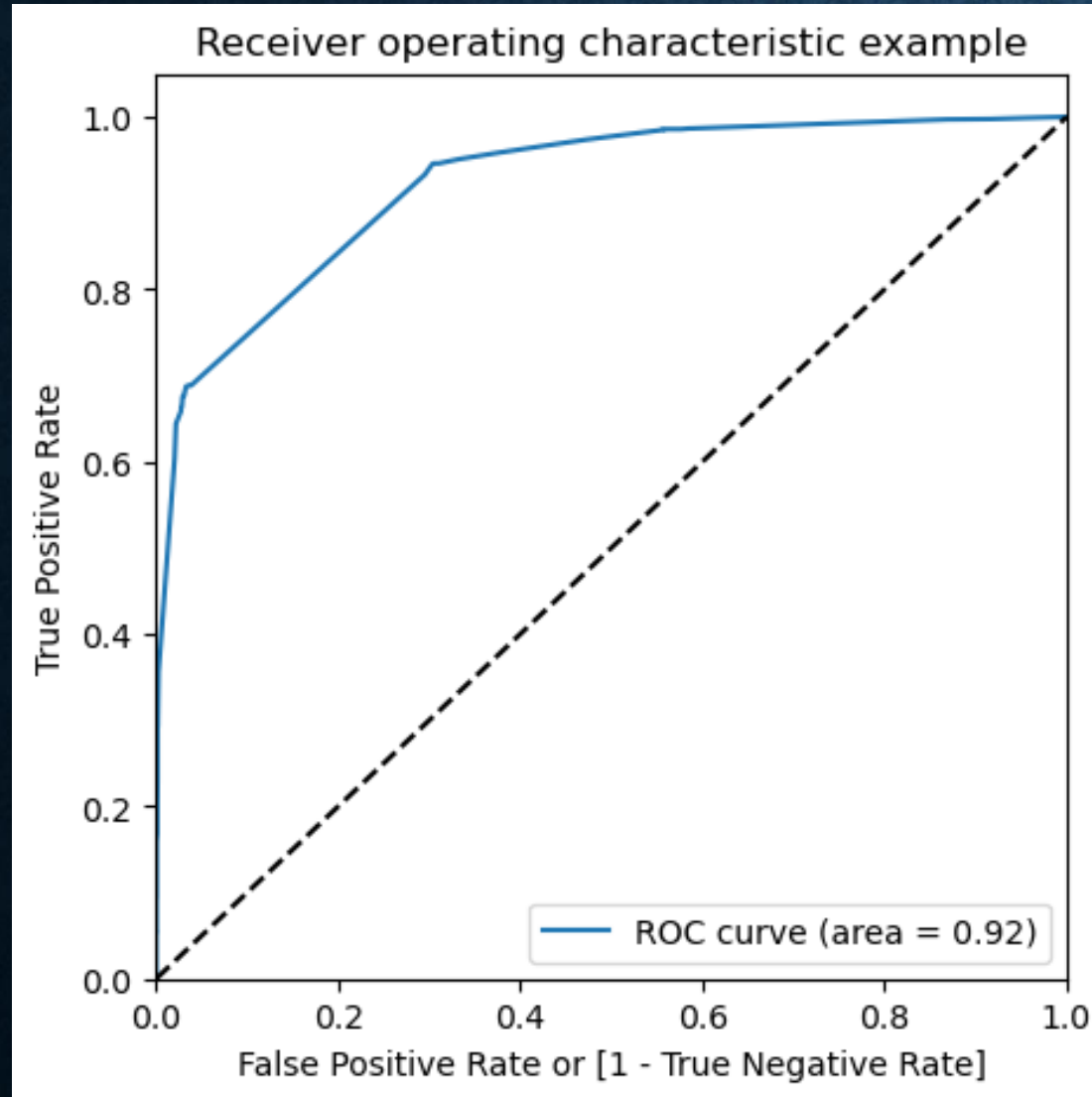
Comparison of leads converted for 'What is your current occupation'



Observation

'Working Professional' has significantly Higher Lead Conversion Rate

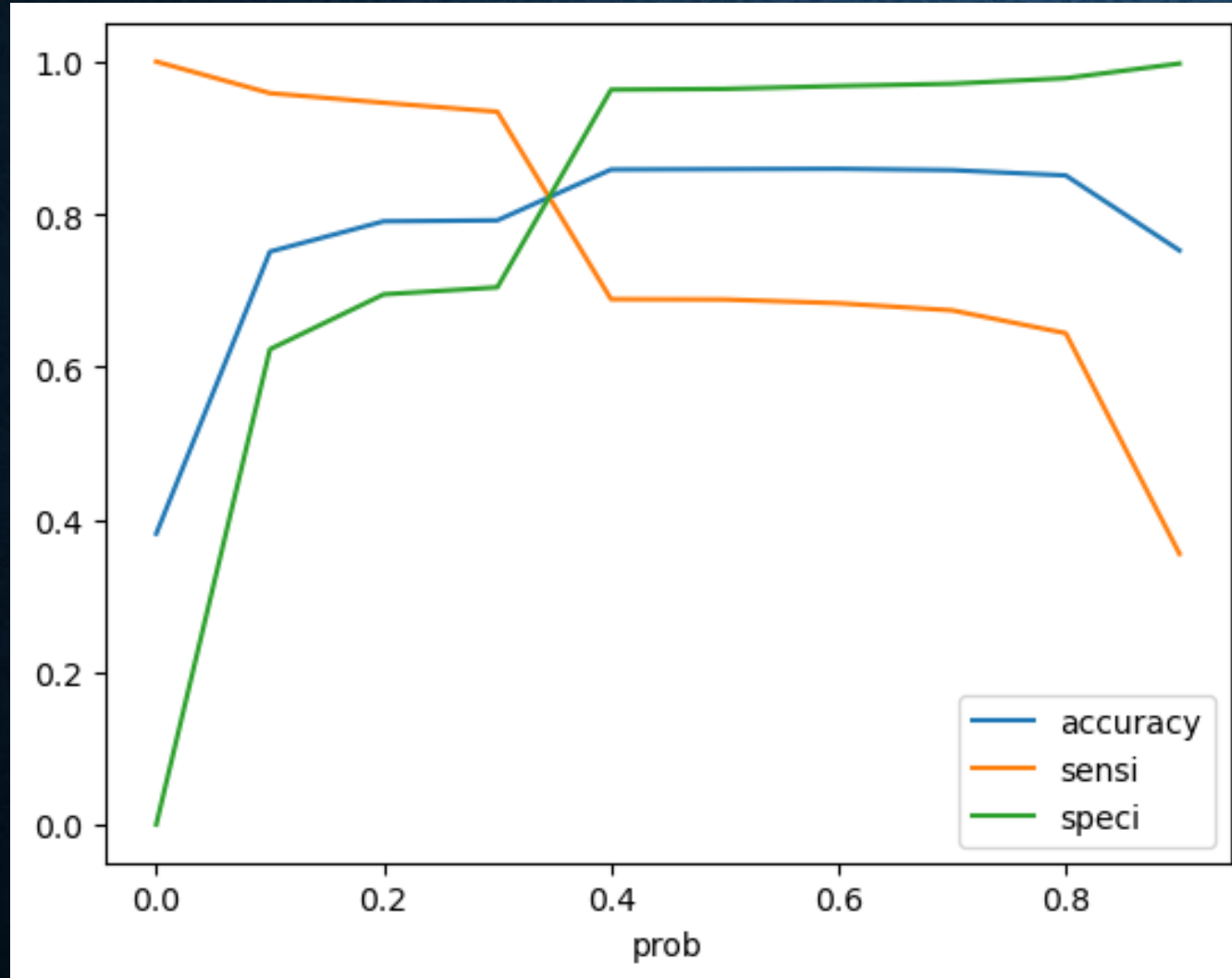
Model Evaluation – 1



Observation

The ROC Curve should be a value close to 1. We are getting good value of 0.92 indicating a good predictive model.

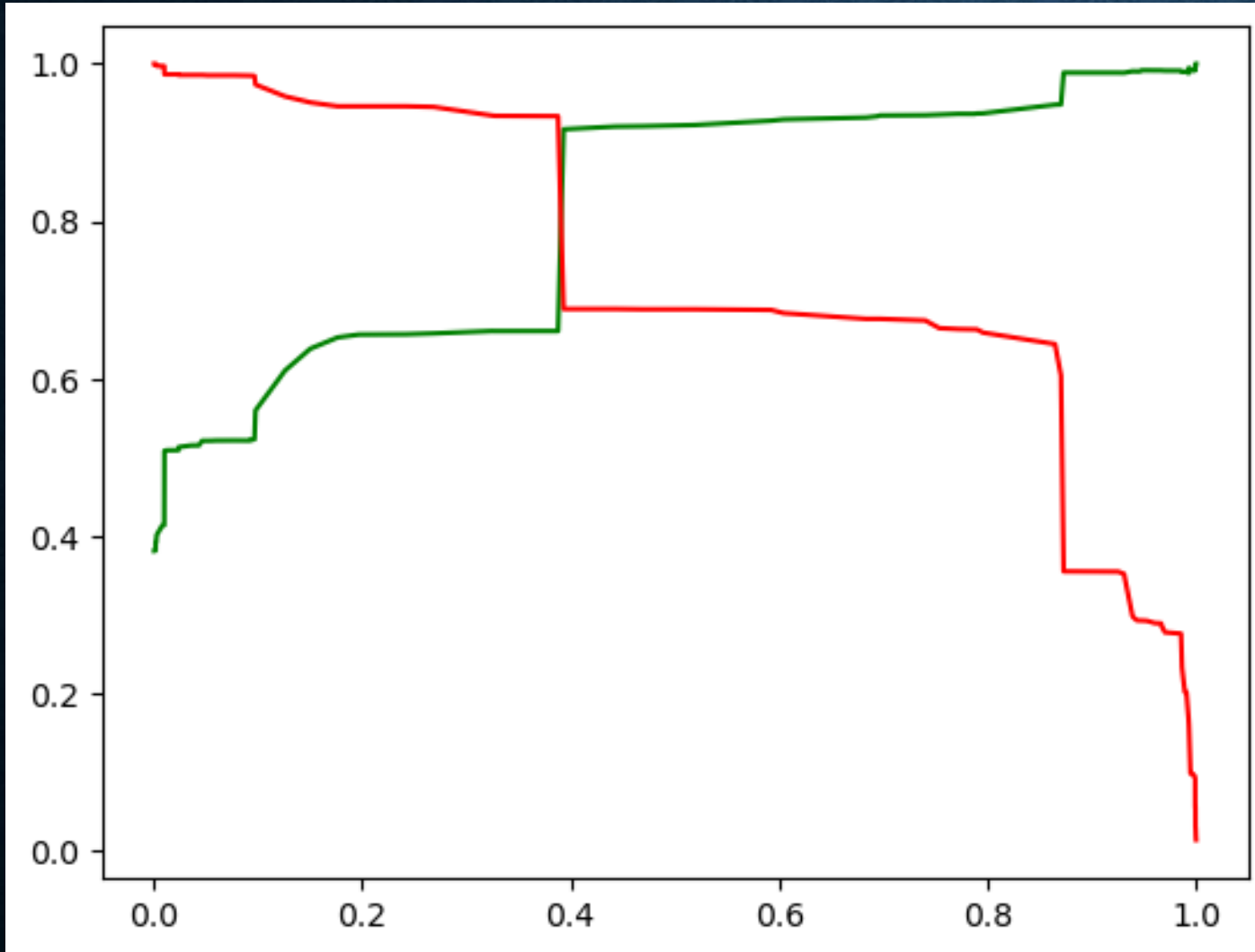
Model Evaluation – 2



Observation

Based on the curve analysis, 0.35 is the optimum point to take it as a cutoff probability.

Model Evaluation – 3



Observation

Based on the curve analysis, the cutoff is 0.38 from above precision and recall chart

Final Observations

Train Data

Accuracy : 79.17%
Sensitivity : 93.34%
Specificity : 70.43%

Test Data

Accuracy : 79.43%
Sensitivity : 93.69%
Specificity : 70.12%

Final Features:

Lead Origin_Lead Add Form
Do Not Email_Yes
Last Activity_Converted to Lead
Last Activity_Olark Chat Conversation
What is your current occupation_Unemployed
What is your current occupation_Working Professional
Tags_Busy
Tags_Closed by Horizzon
Tags_Lost to EINS
Tags_Will revert after reading the email
Tags_in touch with EINS
Last Notable Activity_SMS Sent

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

Conclusion

We know that the relationship between $\ln(\text{odds})$ of 'y' and feature variable "X" is much more intuitive and easier to understand. The equation is:

$$\ln(\text{odds}) = -2.7671 * \text{const} + 2.3606 * \text{Lead Origin_Lead Add Form} - 1.4904 * \text{Do Not Email_Yes} - 1.2735 * \text{Last Activity_Converted to Lead} - 1.7638 * \text{Last Activity_Olark Chat Conversation} - 1.3977 * \text{What is your current occupation_Unemployed} + 1.5888 * \text{What is your current occupation_Working Professional} + 3.1385 * \text{Tags_Busy} + 8.8251 * \text{Tags_Closed by Horizzon} + 8.4480 * \text{Tags_Lost to EINS} - 0.9674 * \text{Tags_Ringing} + 3.7026 * \text{Tags_Will revert after reading the email} + 3.0173 * \text{Tags_in touch with EINS} + 2.4127 * \text{Last Notable Activity_SMS Sent}$$

- Leads which are opening email have high probability to convert thus, Sending SMS & improving customer engagement will also help.
- Working professionals should be targeted as they have High Lead Conversion rate also they have better financials to pay for the course
- Leads whose 'Last Activity' is 'SMS Sent' or 'Olark Chat Conversation' has higher Lead Conversion rate
- Leads with Tags as 'Will revert after reading the email' and 'in touch with EINS' has higher Lead Conversion rate