



Lead Scoring Case Study

Submitted By:

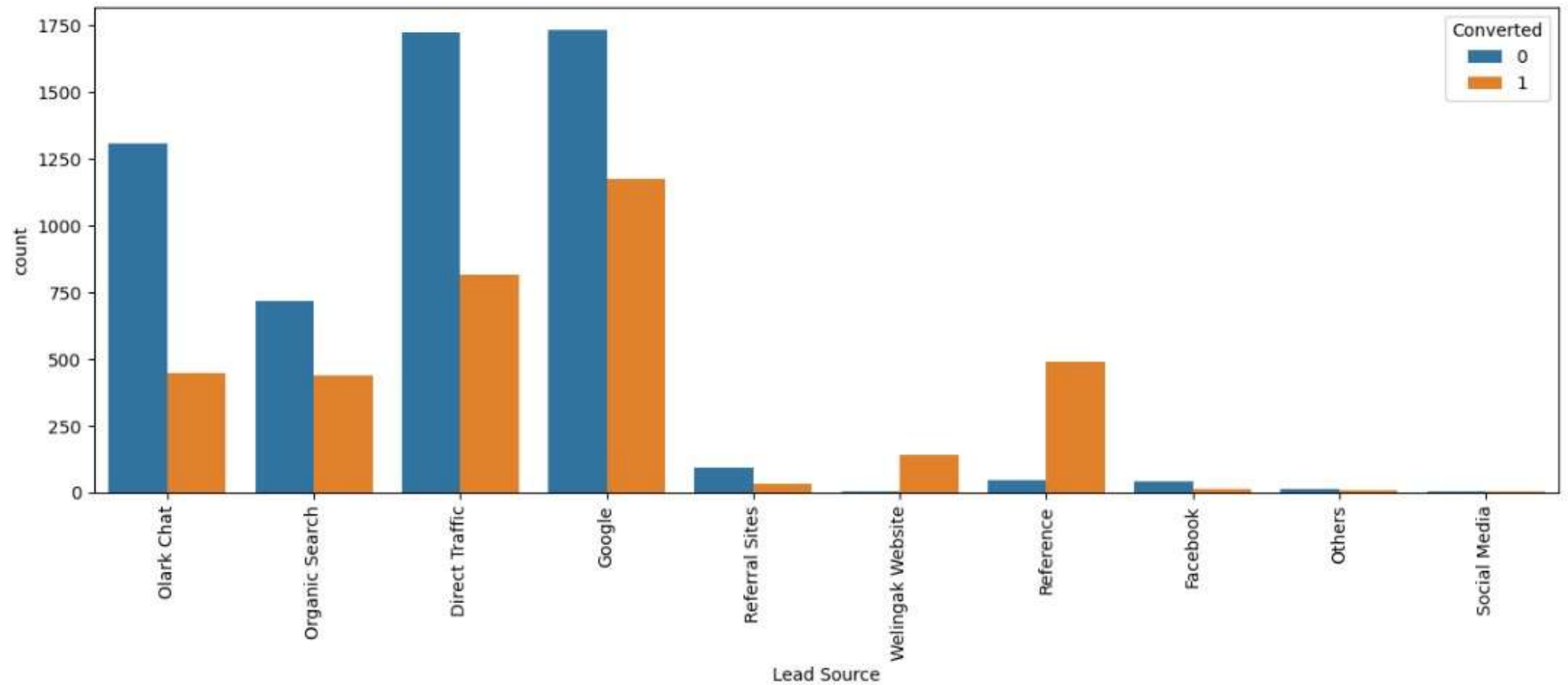
- Harshal Naik
- Meenakshi Naik
- Siddharth Menon

Problem Statement

X Education, an online education company, faces challenges with low lead conversion rates despite generating numerous leads daily. To improve efficiency and focus on potential customers, they seek to identify "Hot Leads" with higher conversion chances. The goal is to build a lead scoring model to assign scores to leads, prioritizing higher conversion chances. By targeting leads with higher scores, the company aims to increase the overall lead conversion rate. The CEO's target is to achieve around 80% lead conversion. To achieve this, we will analyze historical data, including lead source, website interactions, past referrals, and lead characteristics. Through data preprocessing, feature engineering, and machine learning techniques like logistic regression, we will create a lead scoring model. This model will enable the sales team to concentrate efforts on the most promising leads, optimizing lead conversion rates and maximizing the company's revenue.

Goals of the Case Study:

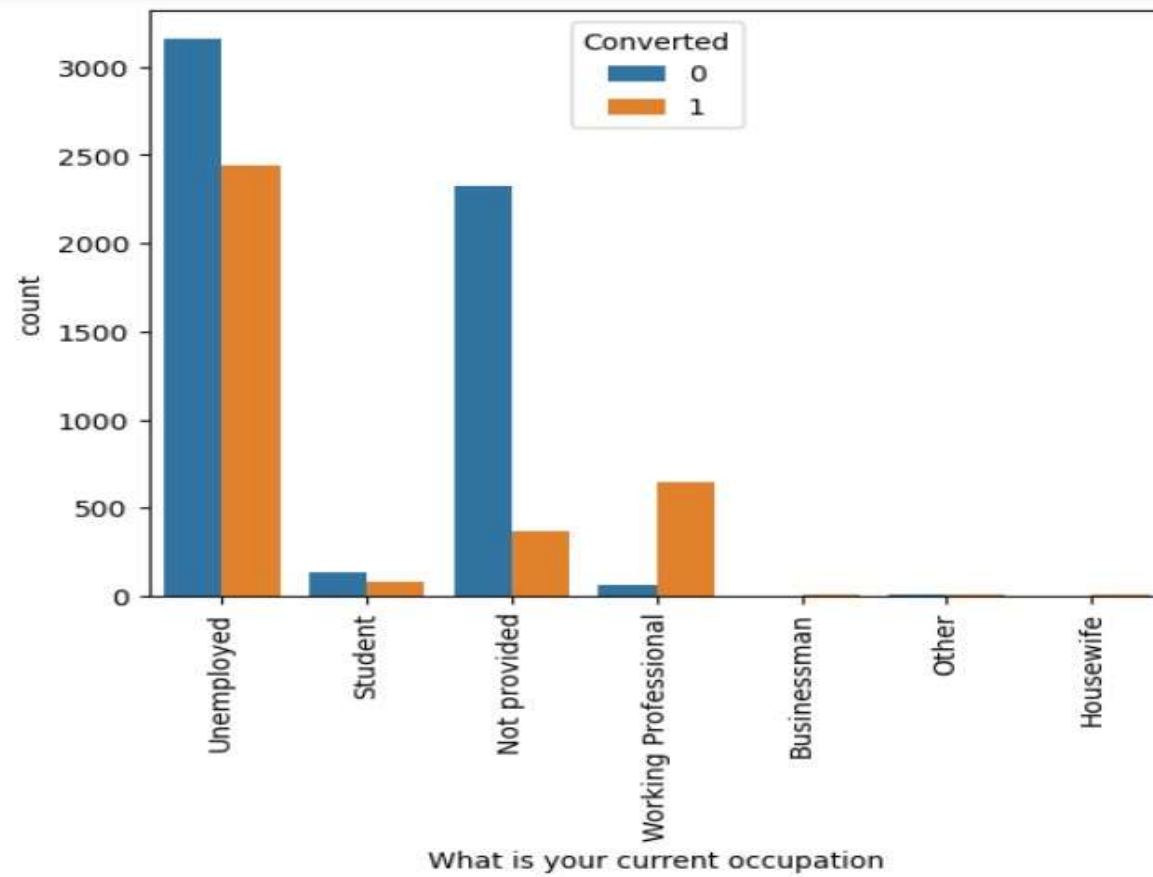
Build a logistic regression model to assign a lead score between 0 and 100 to each of the leads which can be used by the company to target potential leads. A higher score would mean that the lead is hot, i.e., is most likely to convert whereas a lower score would mean that the lead is cold and will mostly not get converted.



Observation

Maximum Leads are generated by Google and Direct Traffic. Conversion rate of Reference leads and Welingak Website leads is very high.

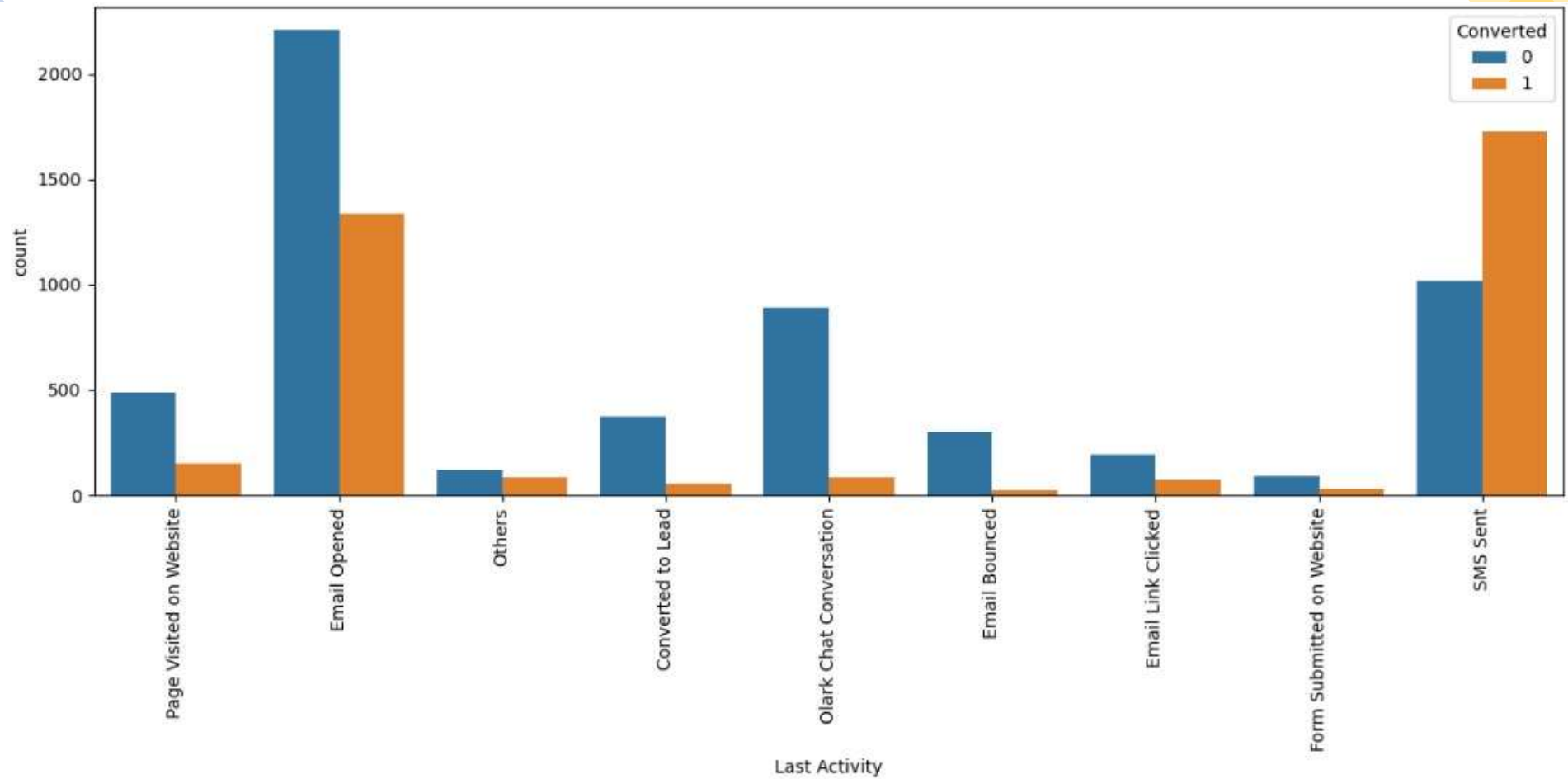
Recommendation: Focus on Reference leads, They are more likely to be converted.



Observation

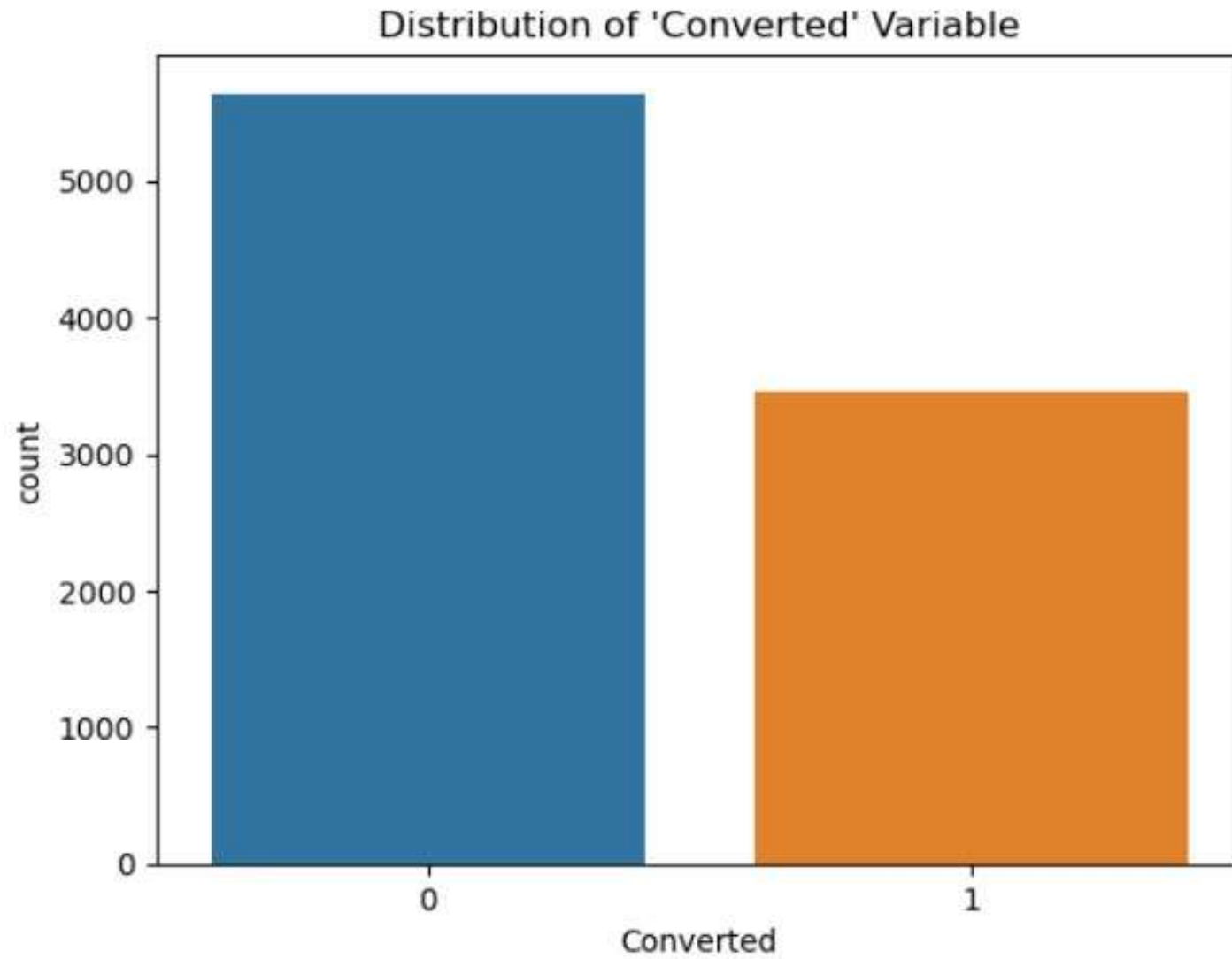
- Maximum leads generated are unemployed and their conversion rate is more than 50%.
- Conversion rate of working professionals is very high.

Working professionals are good potential customers as they are more likely to be financially stable.

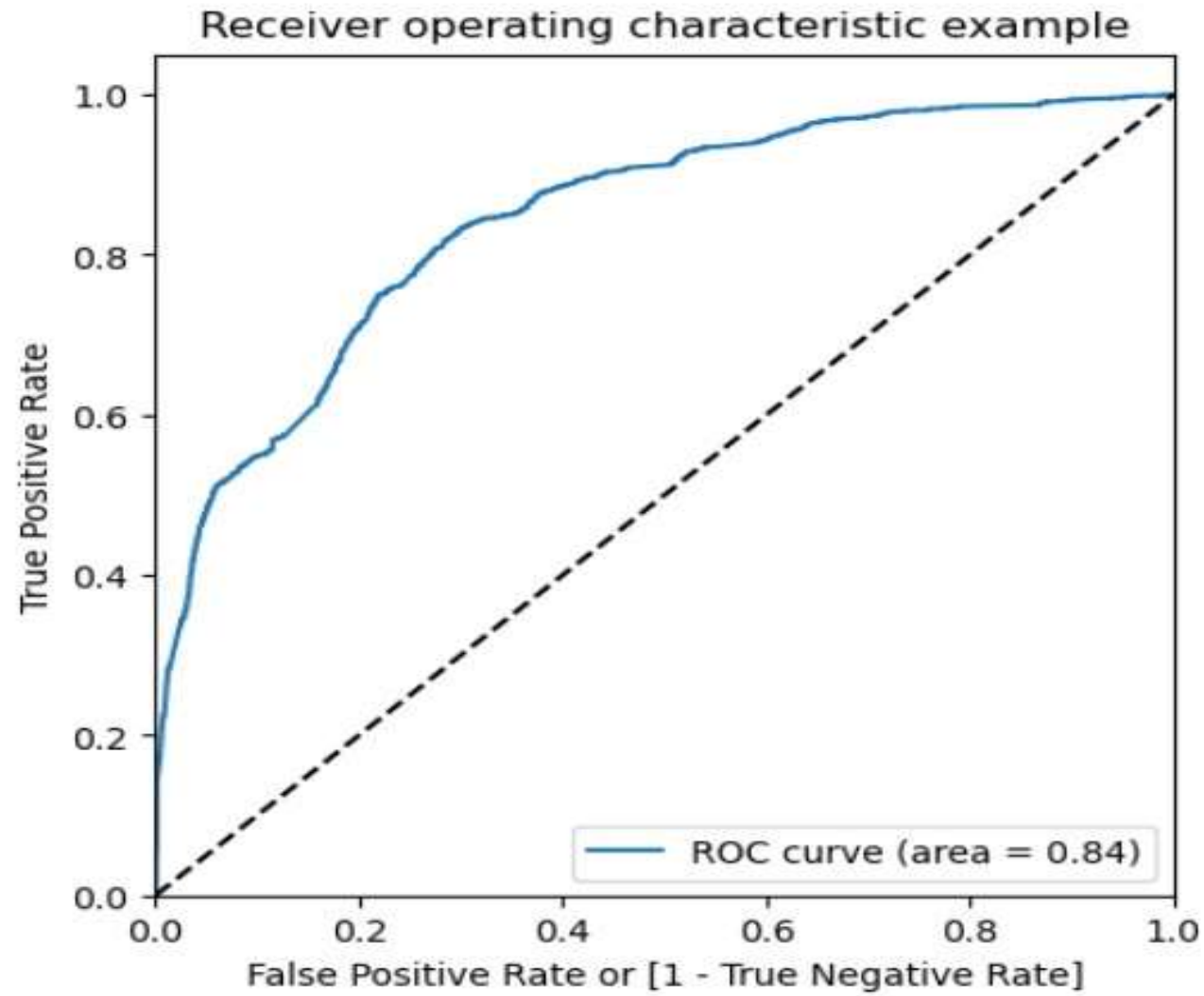


Observation

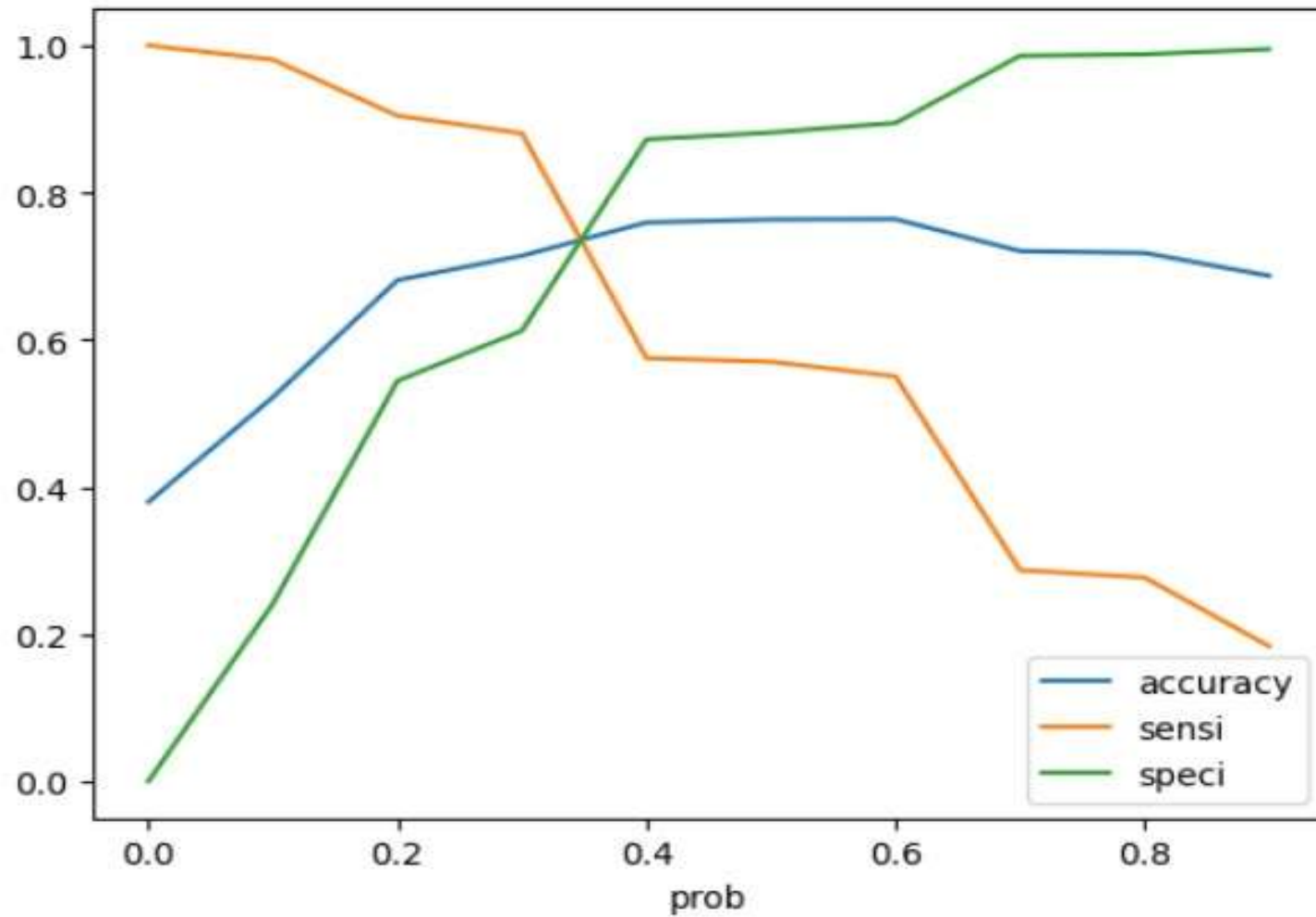
- Maximum leads are generated having last activity as Email opened but conversion rate is not too good.
- SMS sent as last activity has high conversion rate.



Current conversion rate is around 38%



The ROC coefficient is 0.84 indicating accurate analysis by our team



From the curve above, 0.35 is the optimum point to take it as a cutoff probability.

Train and Test Data

| | Train | Test |
|-------------|-------|--------|
| Accuracy | 74% | 74.55% |
| Sensitivity | 84% | 85.22% |
| Specificity | 67% | 67.97% |

Model is performing well on both train and test data.



THANK YOU!