UpGrad

International Institute of Information Technology Bangalore

Lead Score Case Study

Submitted by :

Mr. Muna Sahu

Mr. Pushkar Rajwadikar

UpGrad

International Institute of Information Technology Bangalore

Problem Statement

- X Education offers online courses to industry professionals. Multiple websites and search engines, such as Google, are used by the company to promote its courses.Once these visitors reach the website, they may explore the courses, submit a form for the course, or view videos.
- When these individuals submit a form with their email address or phone number, they are considered leads. Additionally, the organization gets leads from earlier recommendations.
- After acquiring these leads, sales team members begin making calls, sending emails, etc. Through this procedure, some leads are converted, but the majority are not. The average conversion rate for leads at X education is about 30%.

Submitted by :
Mr. Muna Sahu & Mr. Pushkar Rajwadikar
DS C43 | Lead Scoring Case Study | 2022

UpGrad

International Institute of Information Technology Bangalore

Business Goal:

X Education needs help in selecting the most promising leads.

- The leads that are most likely to convert into paying customers
- The company needs a model wherein you a lead score is assigned to each of the leads such that the customers with higher lead score have a higher conversion chance and the customers with lower lead score have a lower conversion chance.
- The CEO, in particular, has given a ballpark of the target lead conversion rate to be around 80%

Submitted by :
Mr. Muna Sahu & Mr. Pushkar Rajwadikar
DS C43 | Lead Scoring Case Study | 2022

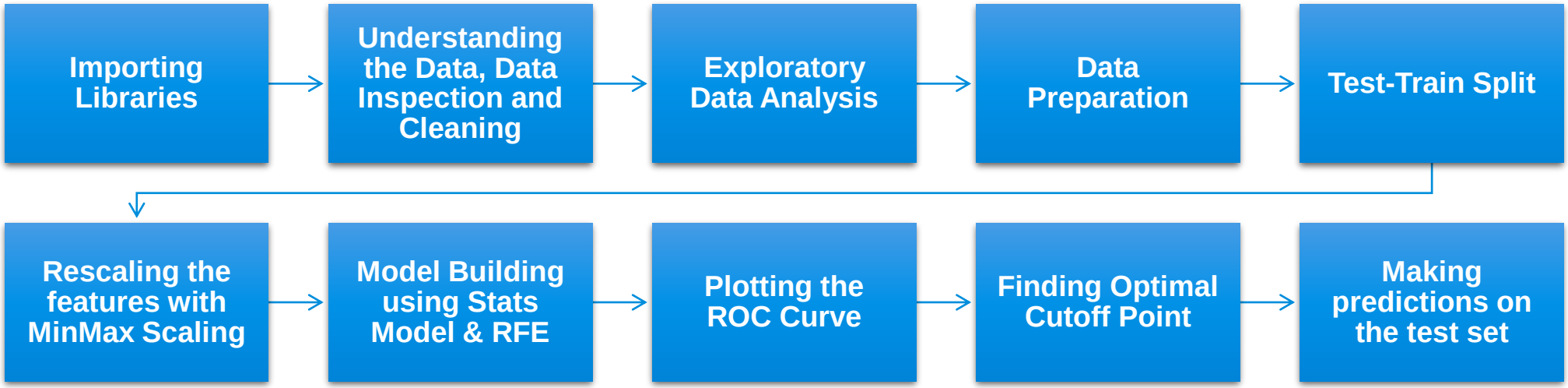
Strategy

- Sourcing the analytical data
- Data cleansing and preparation
- Exploratory data analysis
- Features Scaling
- Splitting the data into Train and Test datasets.
- Building a Logistic Regression model and determine the Lead Score.
- Evaluating the model using several measures, such as Specificity and Sensitivity or Precision and Recall.
- Applying the model with the highest sensitivity and specificity to the Test data.

Submitted by :
Mr. Muna Sahu & Mr. Pushkar Rajwadikar
DS C43 | Lead Scoring Case Study | 2022

Problem solving methodology

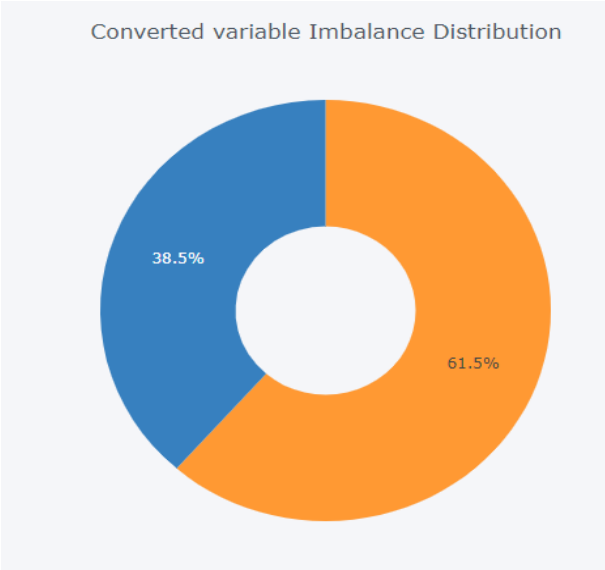
Detailed procedures to solve the business problem

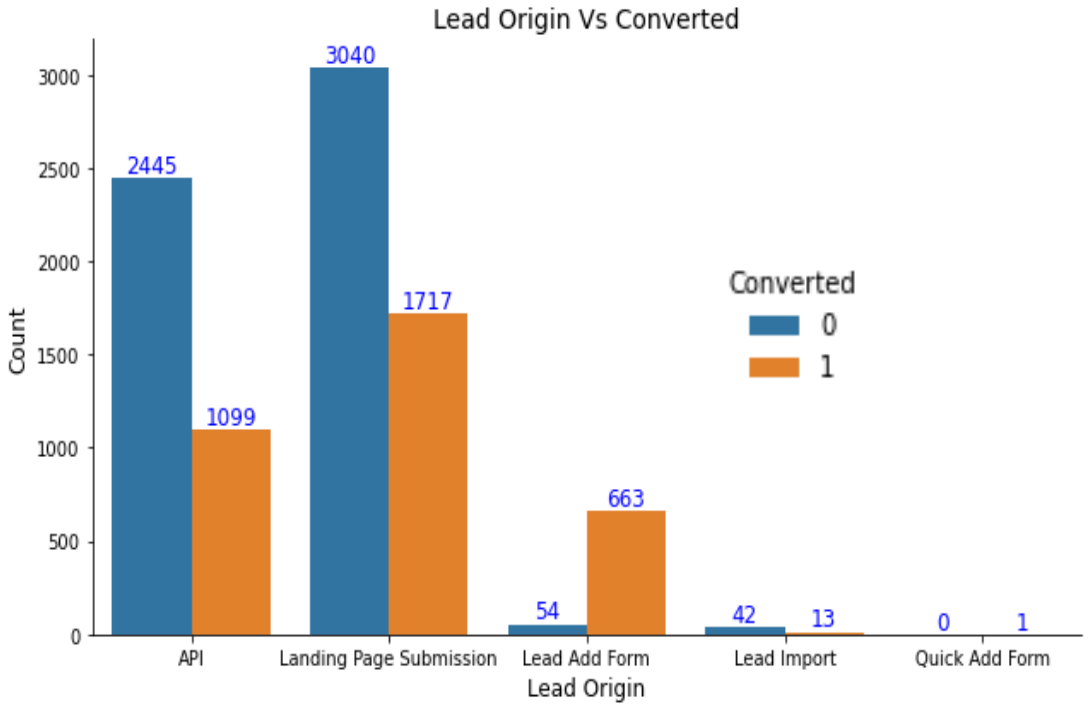


Submitted by :
Mr. Muna Sahu & Mr. Pushkar Rajwadikar

Exploratory Data Analysis

EDA Findings'



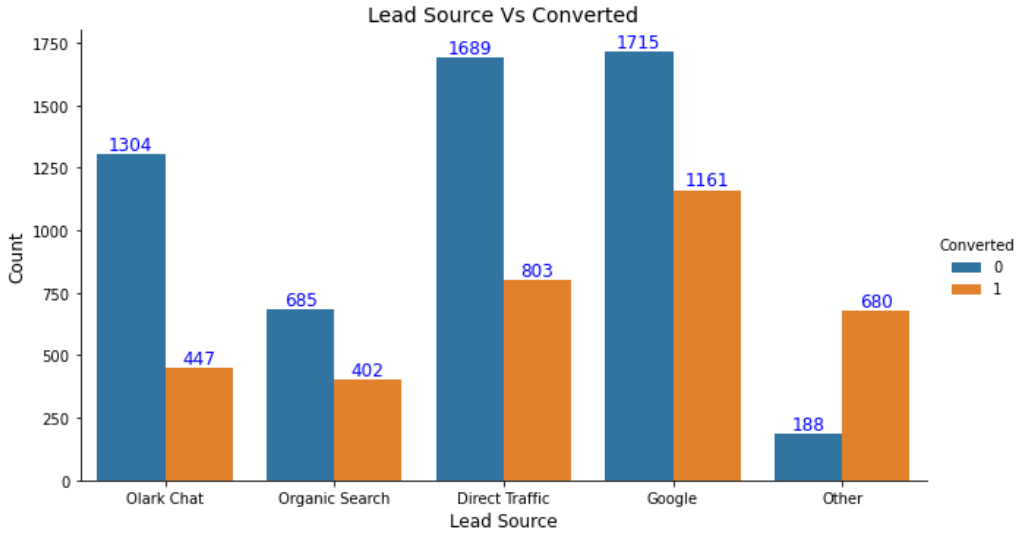


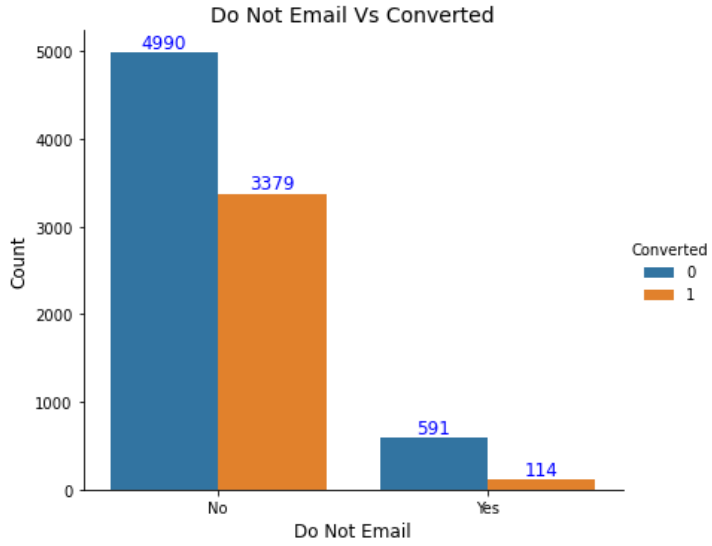
- We have around 39% Conversion rate in Total.
- The conversion rates were high for Total Visits, Total Time Spent on Website and Page Views Per Visit.

Submitted by :
Mr. Muna Sahu & Mr. Pushkar Rajwadikar
DS C43 | Lead Scoring Case Study | 2022

Exploratory Data Analysis

EDA Findings'



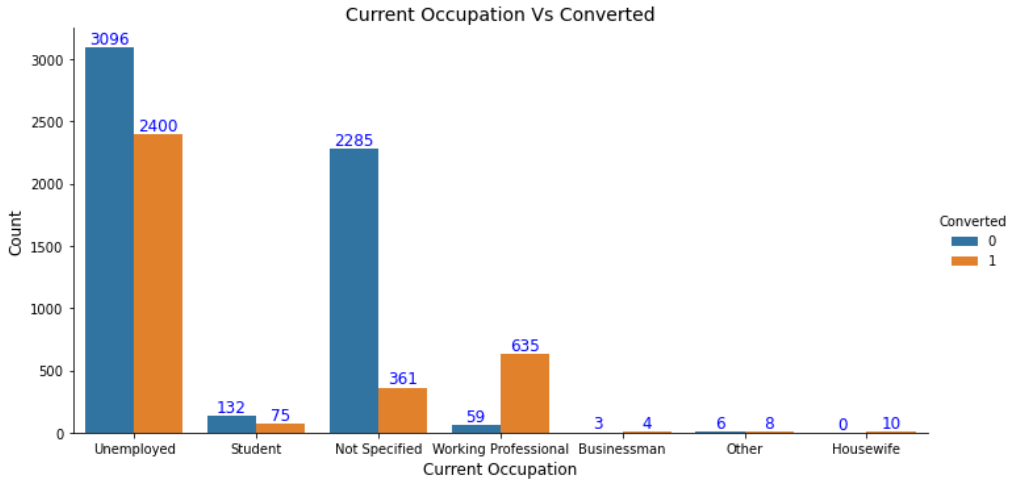


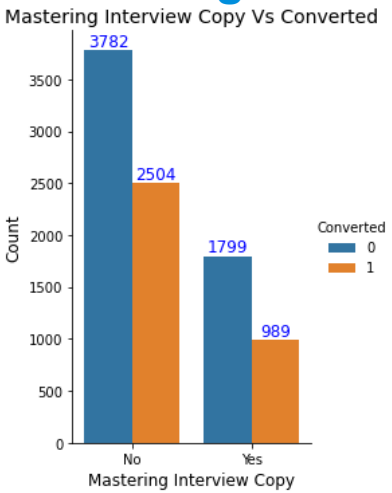
- From above plot we can see that maximum conversions are there when lead is obtained from Google.
- Majority of customers opted to be contacted through emails and the conversion rate is also higher for these customers. Also, conversion rate is very for customers opting to not receive email from the company.

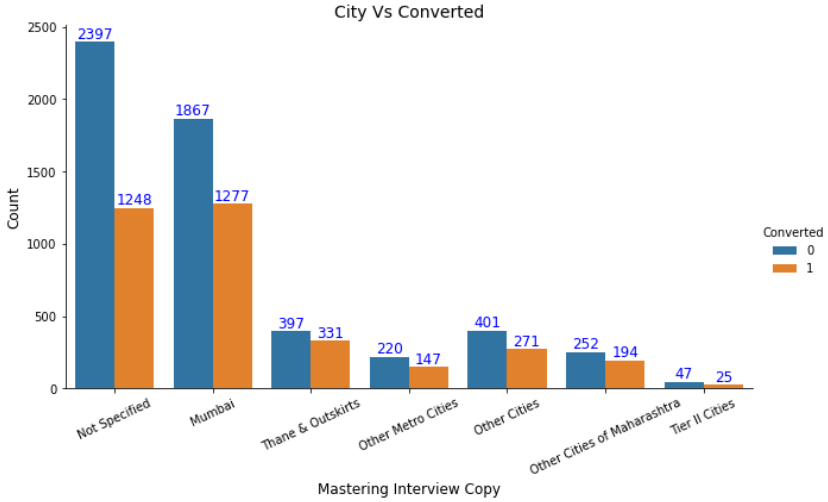
Submitted by :
Mr. Muna Sahu & Mr. Pushkar Rajwadikar
DS C43 | Lead Scoring Case Study | 2022

Exploratory Data Analysis

EDA Findings'





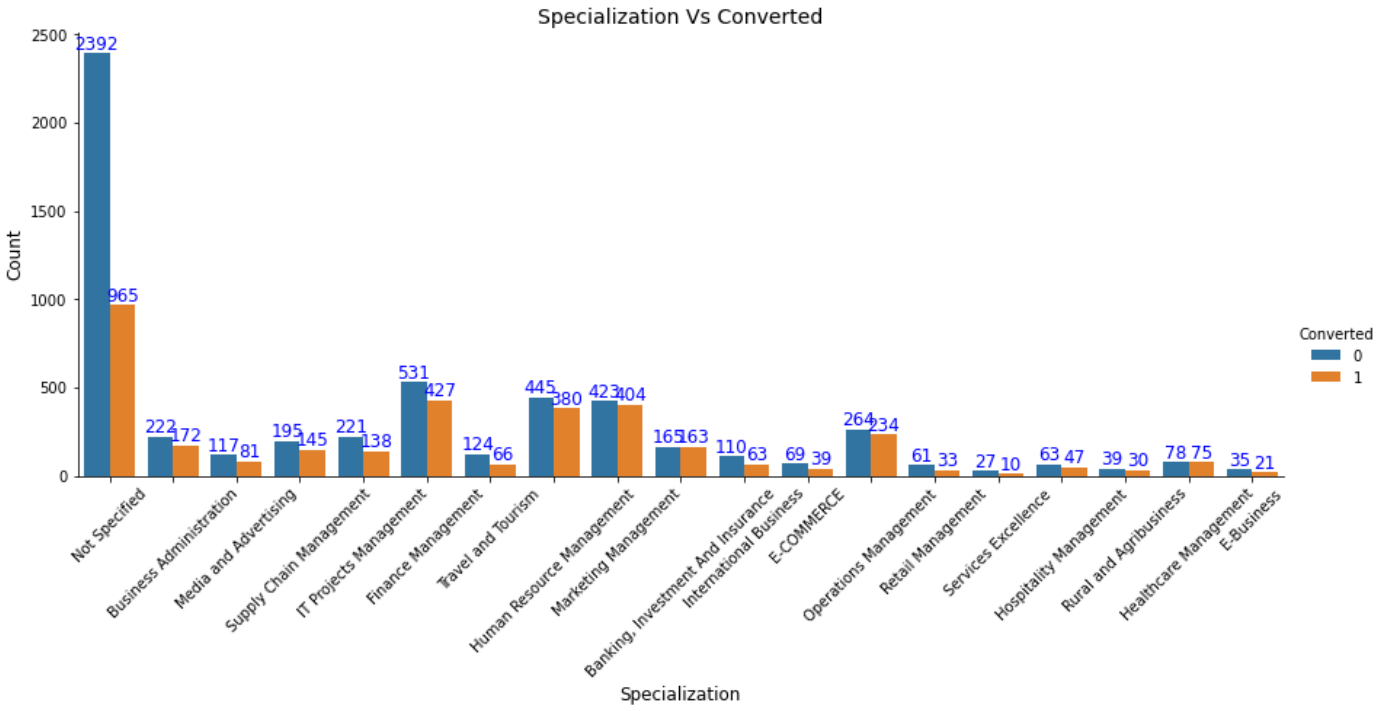


- Maximum customers that converted are unemployed. We can also observe that working professionals are more likely to convert where customers who chose not to share their occupation are less likely to be converted.
- Customers who do not want a free copy of Mastering the Interview have higher conversion rate.
- The graph shows that maximum number of leads are from Mumbai whereas conversion rate is higher for Thane Outskirts.

Submitted by :
Mr. Muna Sahu & Mr. Pushkar Rajwadikar
DS C43 | Lead Scoring Case Study | 2022

Exploratory Data Analysis

EDA Findings'



- Though majority leads are the customers who have not specified their Specialization, the conversion rate is very low.
- We can also see that lead conversion rate is higher for Management specializations (Finance, marketing, operations) as well as Banking & Insurance.

Submitted by :
Mr. Muna Sahu & Mr. Pushkar Rajwadikar
DS C43 | Lead Scoring Case Study | 2022

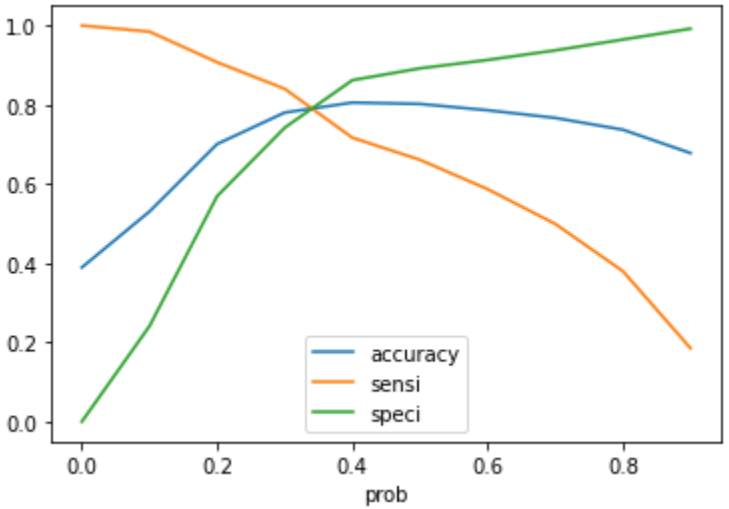
Variables Impacting the Conversion Rate

1. Total Time Spent on Website
2. Lead Origin_Lead Add Form
3. Occupation_Working Professional
4. Occupation_Other
5. Occupation_Unemployed
6. Lead Source_Olark Chat
7. Occupation_Student
8. TotalVisits
9. Do Not Email

Submitted by :
Mr. Muna Sahu & Mr. Pushkar Rajwadikar

Model Evaluation

Sensitivity and Specificity on Train Data Set



accuracy sensitivity and specificity for various probabilities.

2197	962
495	2067

Confusion Matrix

- From the curve above, 0.31 is the optimum point to take it as a cutoff probability.

Accuracy : 78 %

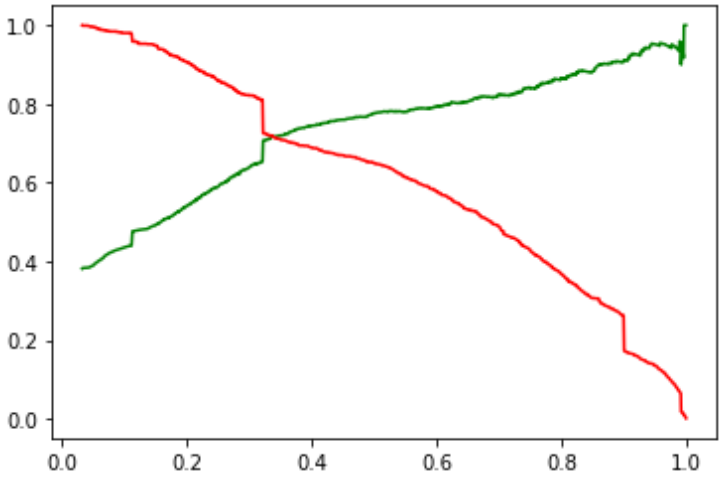
Sensitivity: 84 %

Specificity: 75 %

Submitted by :
Mr. Muna Sahu & Mr. Pushkar Rajwadikar

Model Evaluation

Sensitivity and Specificity on Test Data Set



accuracy sensitivity and specificity for various probabilities.

1247	455
184	837

Confusion Matrix

- From the curve above, 0.31 is the optimum point to take it as a cutoff probability.

Accuracy : 77 %

Sensitivity: 82 %

Specificity: 73 %

Submitted by :
Mr. Muna Sahu & Mr. Pushkar Rajwadikar

THANK YOU!

UpGrad

International Institute of Information Technology Bangalore

Mr Muna Sahu

DS C43 | 2022

+91 9348327393

munasahoo012@gmail.com

Mr. Pushkar Rajwadikar

DS C43 | 2022

+91 7057855533

pushkarrajwadikar2014@gmail.com