



Lead Scoring- Case Study

By Sougata Saha

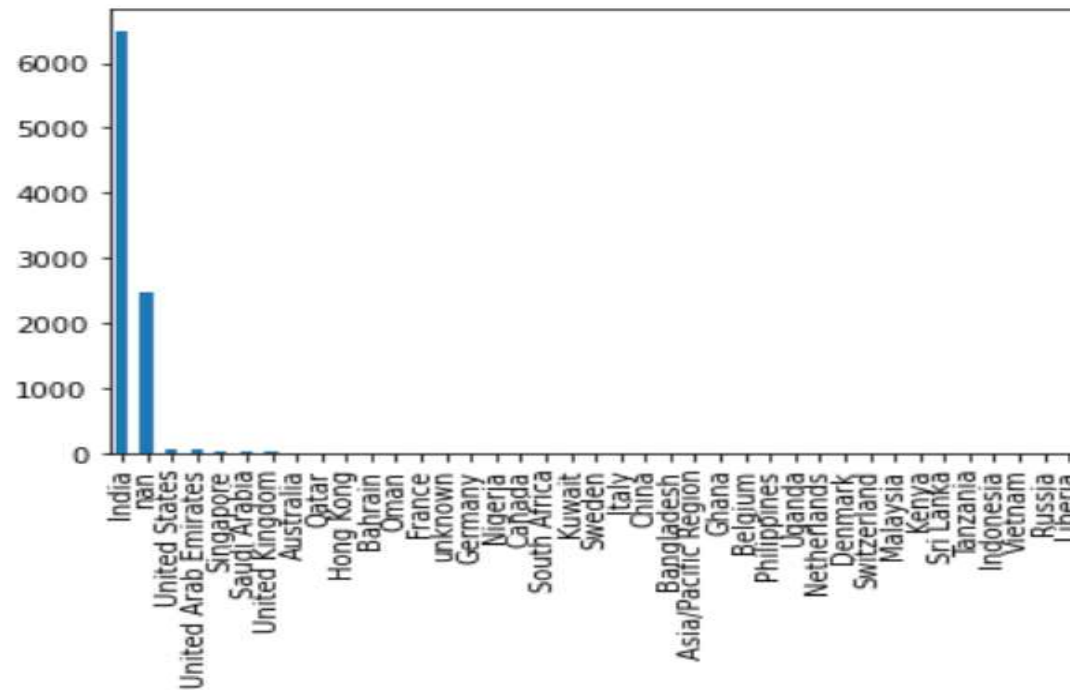


Business Problem

- ▶ The X Education company wants to identify the hot leads, who can be targeted to sell their courses
- ▶ Currently, the company is able to convert only 30% of the total leads to take up the courses in their portal

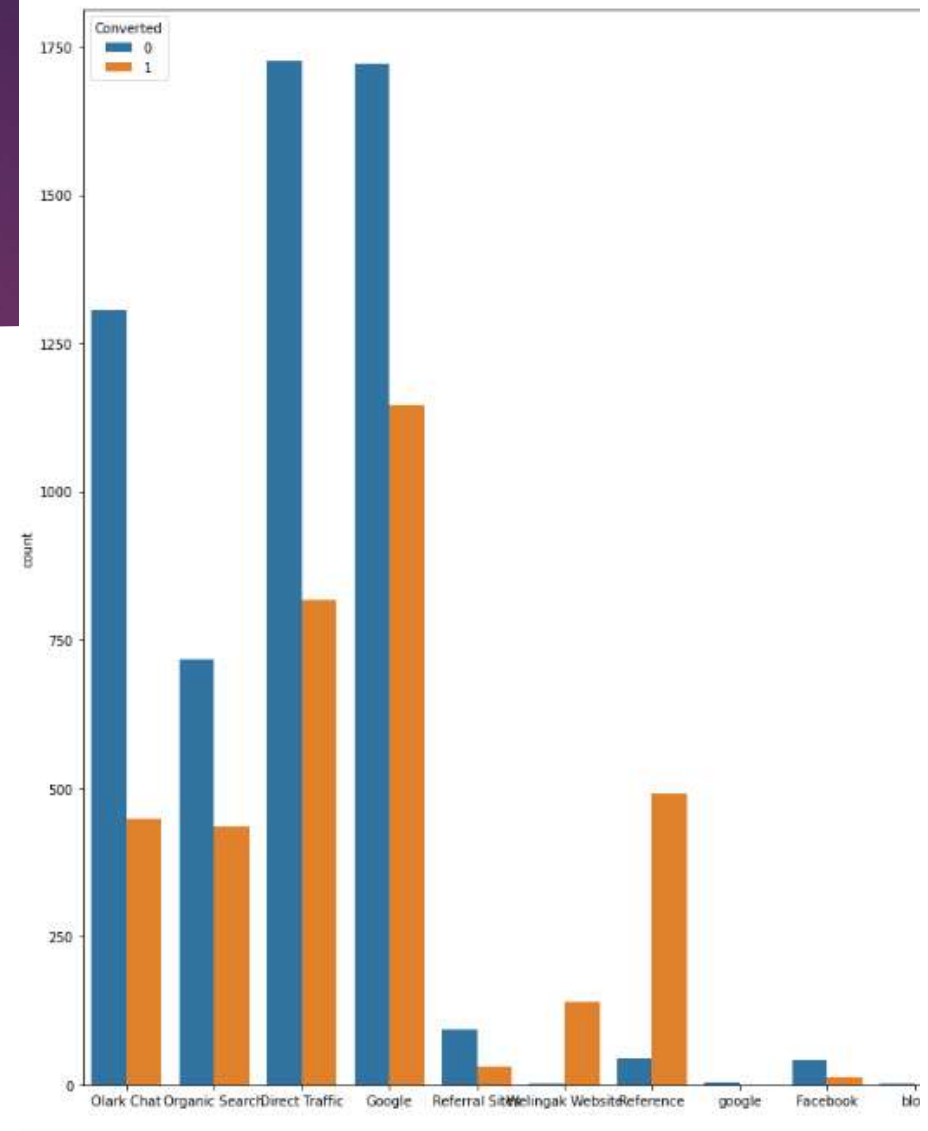
Initial Insights from Data

- ▶ India is predominately the country that shows interest in this course, hence that has been dropped for other analysis. Similary features with maximum number of single values are dropped



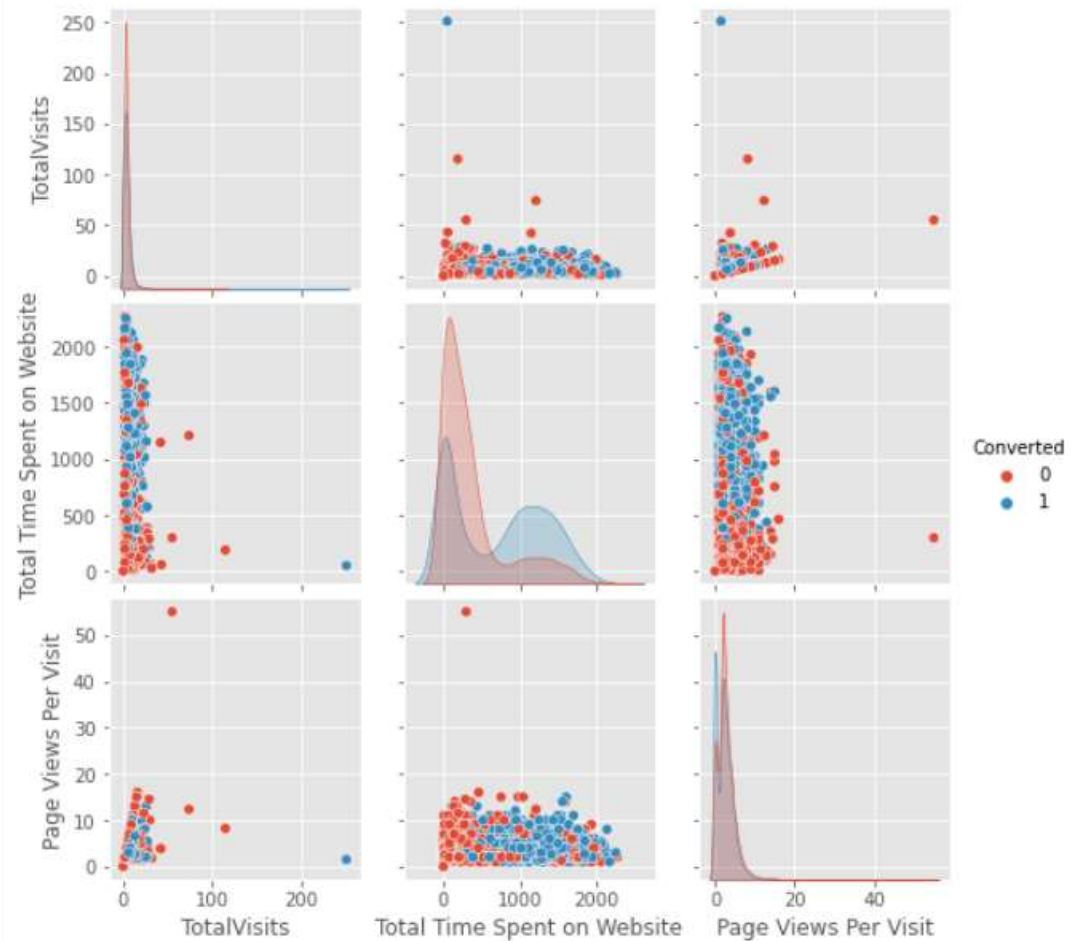
EDA

- ▶ In terms of lead source, Google and Direct traffic are preferred.
- ▶ But the conversion rate is higher in Google and Olark chat



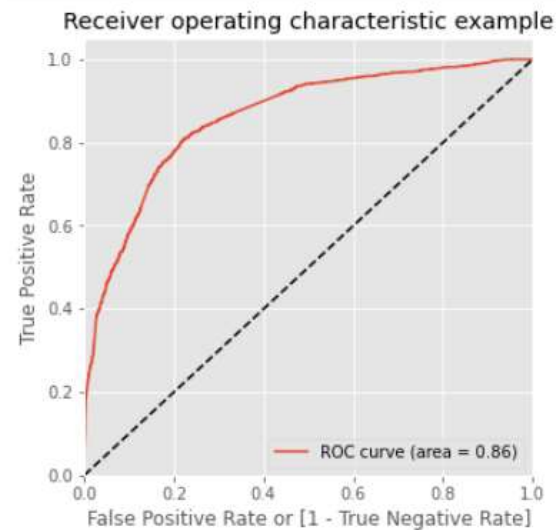
EDA

- It can be seen from this pair plot, that the total number of visits and total time spent increases the conversion rate also increase.



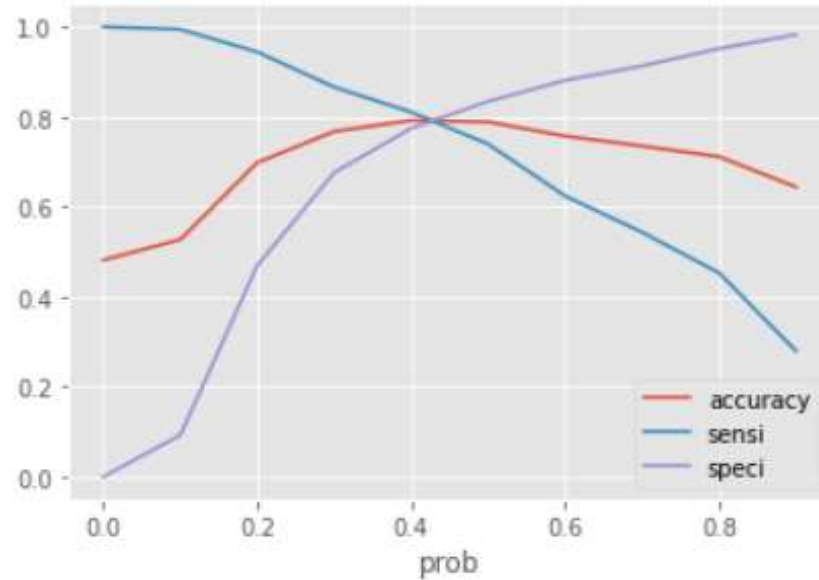
ML model

- ROC curve if away from right angle line, the model built would be efficient. The ROC curve of 0.86 has been achieved implying it a good model.



Optimal Cut-off

- To do away with the trade-off between the various metrics in ML model, it's prudent to have an optimal cut-off value. In this model, the value is 0.42



Recommendations

- The total time spent and more number of visits to website play a crucial role in getting

	coef	std err	z	P> z	[0.025	0.975]
const	0.2040	0.196	1.043	0.297	-0.179	0.587
TotalVisits	11.1489	2.665	4.184	0.000	5.926	16.371
Total Time Spent on Website	4.4223	0.185	23.899	0.000	4.060	4.785
Lead Origin_Lead Add Form	4.2051	0.258	16.275	0.000	3.699	4.712
Lead Source_Olark Chat	1.4526	0.122	11.934	0.000	1.214	1.691
Lead Source_Welingak Website	2.1526	1.037	2.076	0.038	0.121	4.185
Do Not Email_Yes	-1.5037	0.193	-7.774	0.000	-1.883	-1.125
Last Activity_Had a Phone Conversation	2.7552	0.802	3.438	0.001	1.184	4.326
Last Activity_SMS Sent	1.1856	0.082	14.421	0.000	1.024	1.347
What is your current occupation_Student	-2.3578	0.281	-8.392	0.000	-2.908	-1.807
What is your current occupation_Unemployed	-2.5445	0.186	-13.699	0.000	-2.908	-2.180
Last Notable Activity_Unreachable	2.7846	0.807	3.449	0.001	1.202	4.367

Recommendation

- ▶ Those who have opted for Do No Email on the portal have less interested in getting into the course □ Interestingly, phone conversation shows a higher conversion rate
- ▶ Wellingak's website shows a higher coefficient for getting enrolled to the course.
- ▶ If the lead origin is Lead Add Form, the chances for buyers purchasing this course is very high compared to other means.

Metrics evaluation

Metrics of the train data before using optimal cut-off

```
In [ ]: 1 Accuracy: 78%  
        2 Specificity: 83%  
        3 Sensitivity : 73%
```

Metrics of the train data after using optimal cut-off

```
In [ ]: 1 Accuracy : 79%  
        2 Specificity: 79%  
        3 Sensitivity: 79%
```

Metrics of the test data

```
In [ ]: 1 Accuracy : 79%  
        2 Specificity: 80%  
        3 Sensitivity: 76%
```