

LEAD SCORING CASE STUDY

SUBMITTED BY -Ruchira Dolai






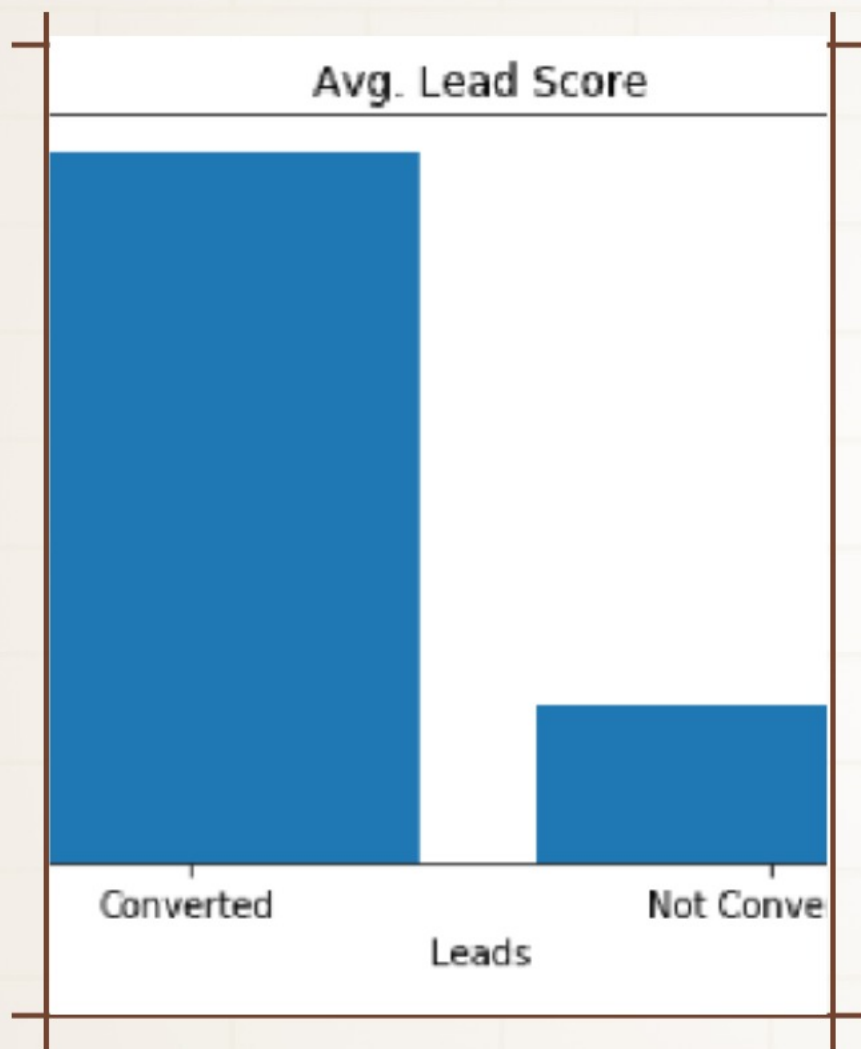
- ★ Objective:
X Education expects to help them select the most promising leads, i.e. the leads that are most likely to convert into paying customers. The company wants us to build a model wherein we need to assign a lead score to each of the leads such that the customers with higher lead score have 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%.

ANALYSIS approach from dataset

Data
understanding
and Exploring
From leads
dataset

Prospect ID	# Lead Number	Lead Origin	Lead Source	
9240 unique values		Landing Page Sub... 53%	Google 31%	
		API 39%	Direct Traffic 28%	
		Other (774) 8%	Other (3829) 41%	
7927b2df-8bba-4d29-b9a2-b6e0beafe620	660737	API	Olark Chat	No
2a272436-5132-4136-86fa-dcc88c88f482	660728	API	Organic Search	No
8cc8c611-a219-4f35-ad23-fdfd2656bd8a	660727	Landing Page Submission	Direct Traffic	No
0cc2df48-7cf4-4e39-9de9-19797f9b38cc	660719	Landing Page Submission	Direct Traffic	No
3256f628-e534-4826-9d63-4a8b88782852	660681	Landing Page Submission	Google	No
2058ef08-2858-443e-a01f-a9237db2f5ce	660680	API	Olark Chat	No
9fae7dfa-169d-489b-	660673	Landing Page	Google	No

Data Cleaning and
Prepare for data
model



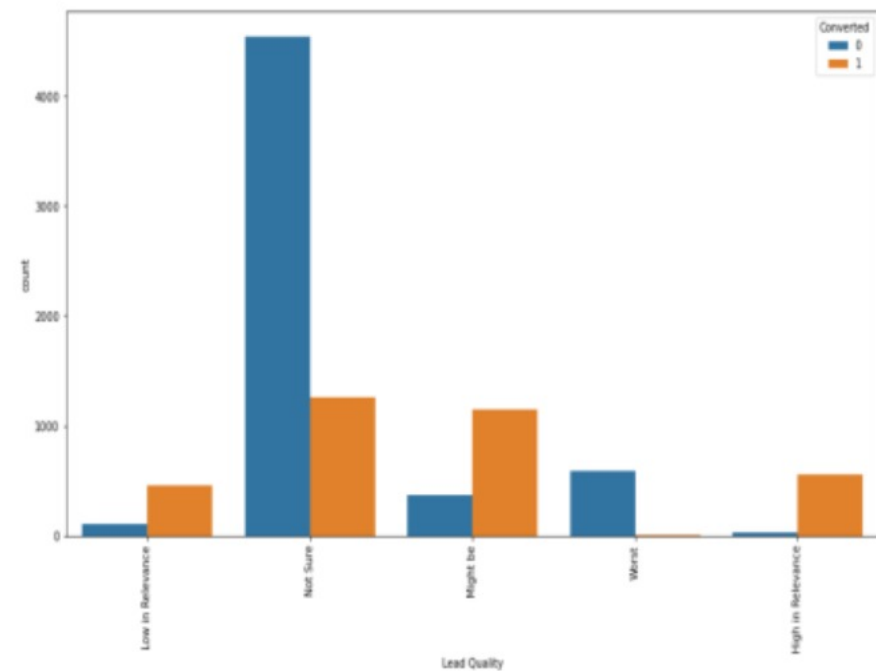
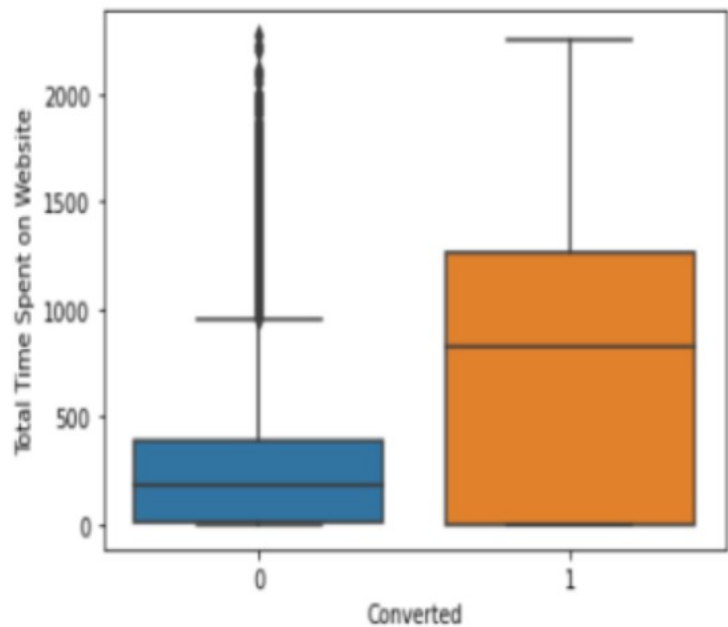
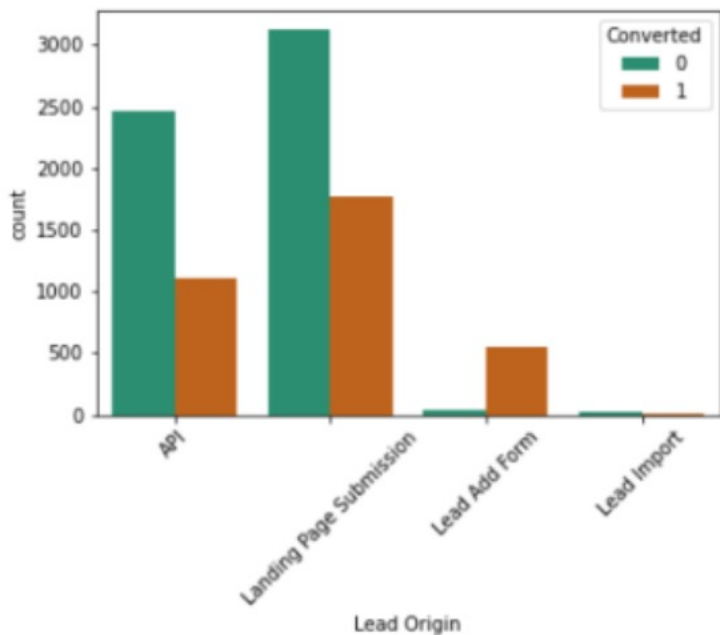
Model outcome

Model outcome for lead score above 35.

Higher the lead score, more chance the lead customer to get converted.

Average lead score of converted leads 68, of non converted leads 15.

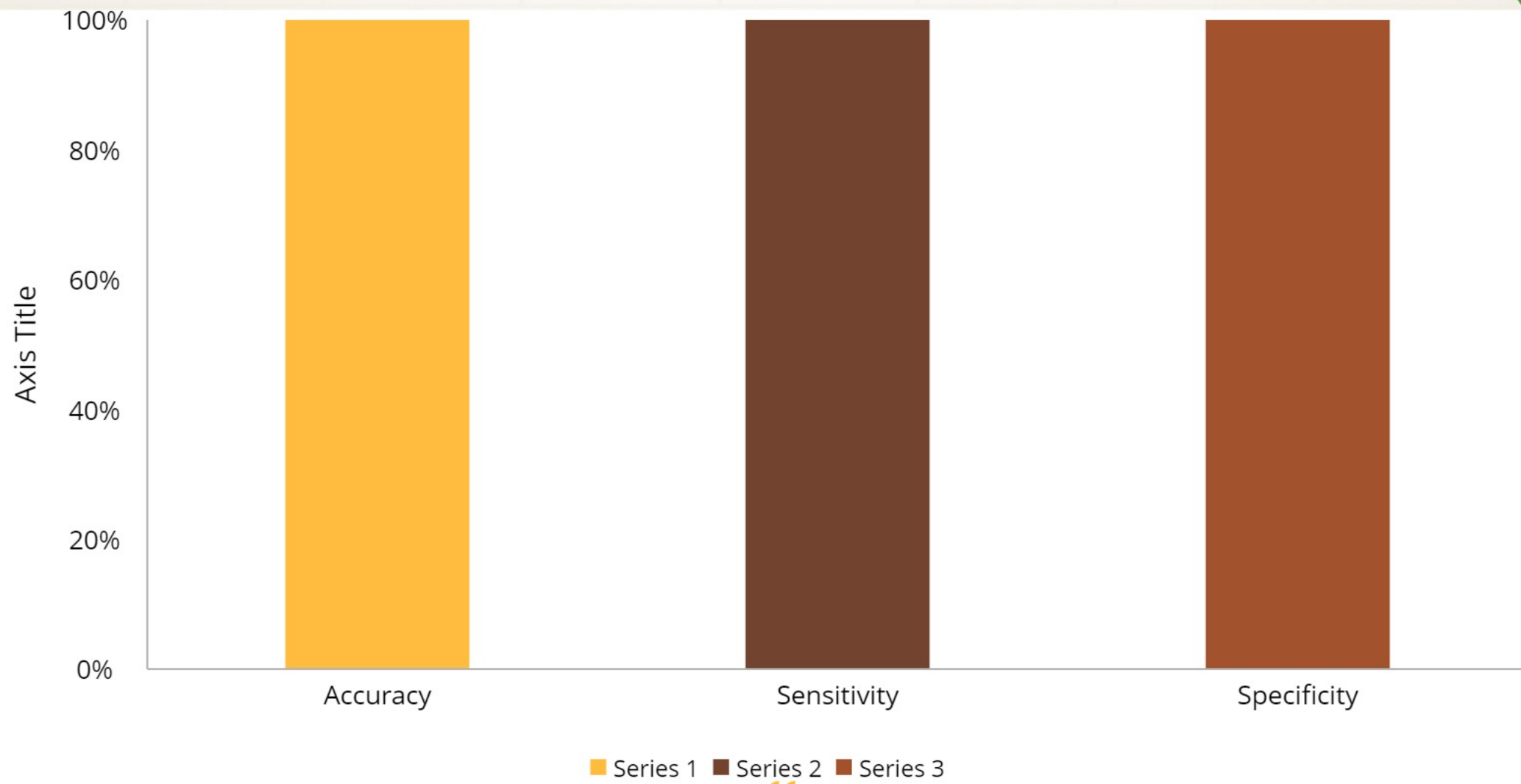
Optimum probability cutoff is 0.5



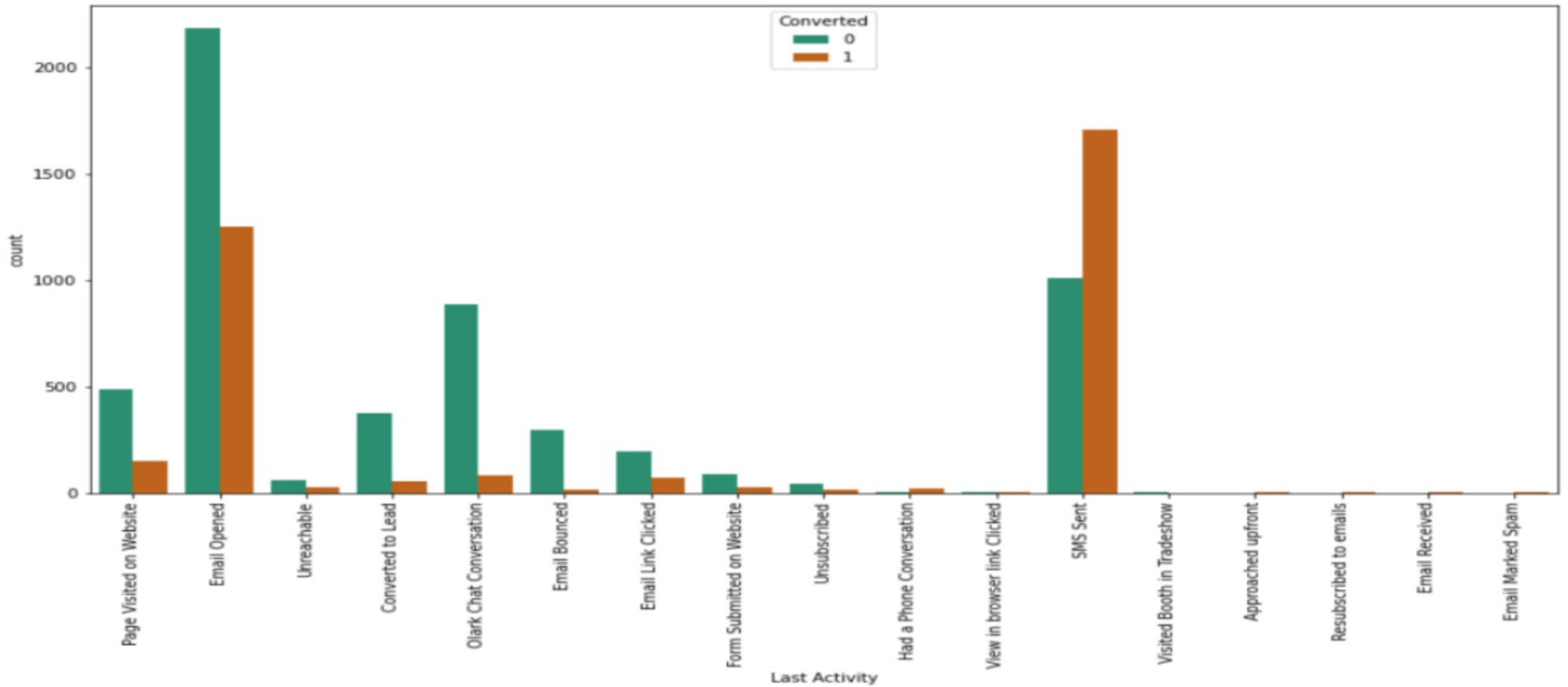
Lead Origin: Lead Add form highest conversion.

Total time spent on website: People with most total time have high conversion rate.

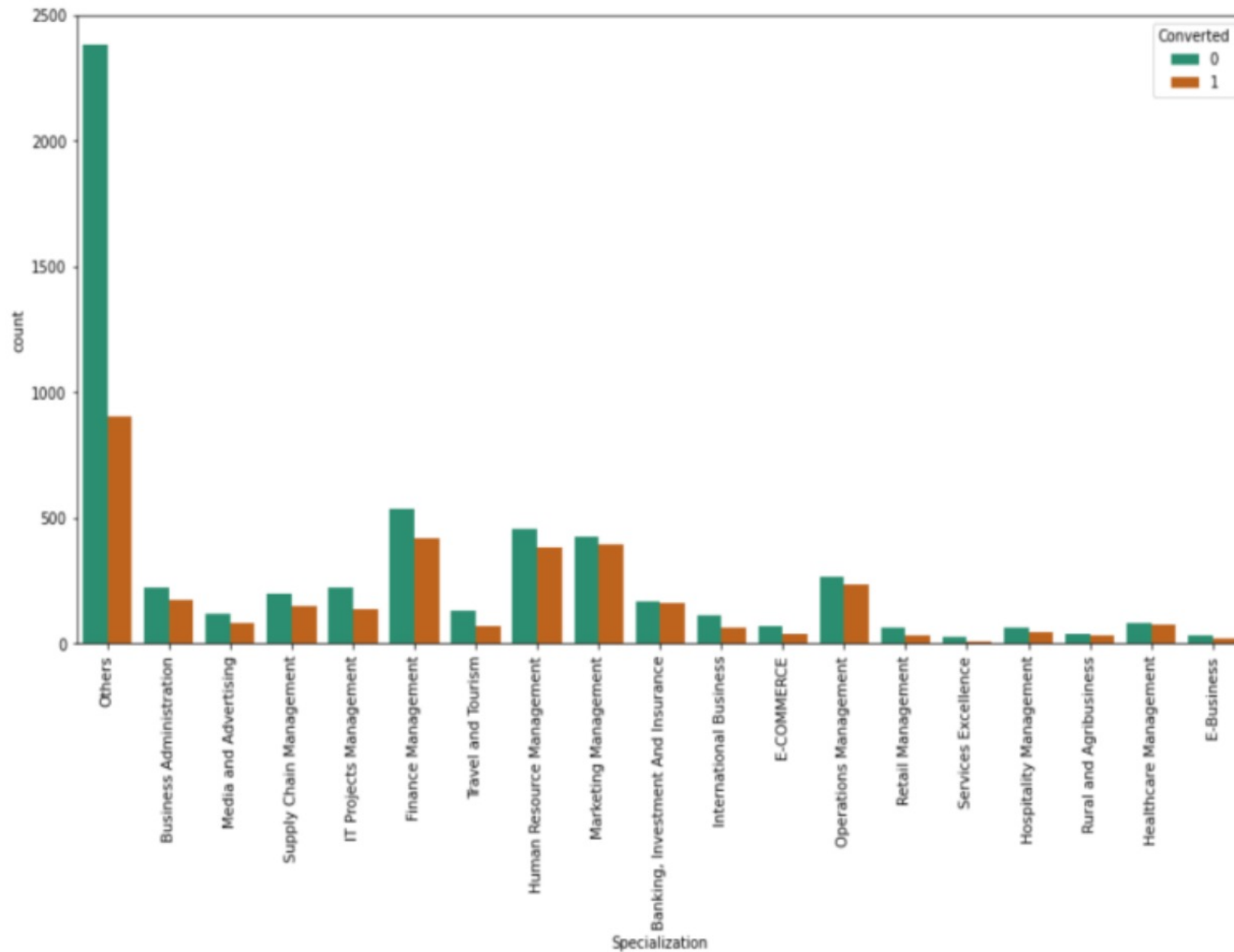
Lead Quality: Efforts to be made to correct idea of relevance of person incharge of lead.



Accuracy Specificity and Sensitivity

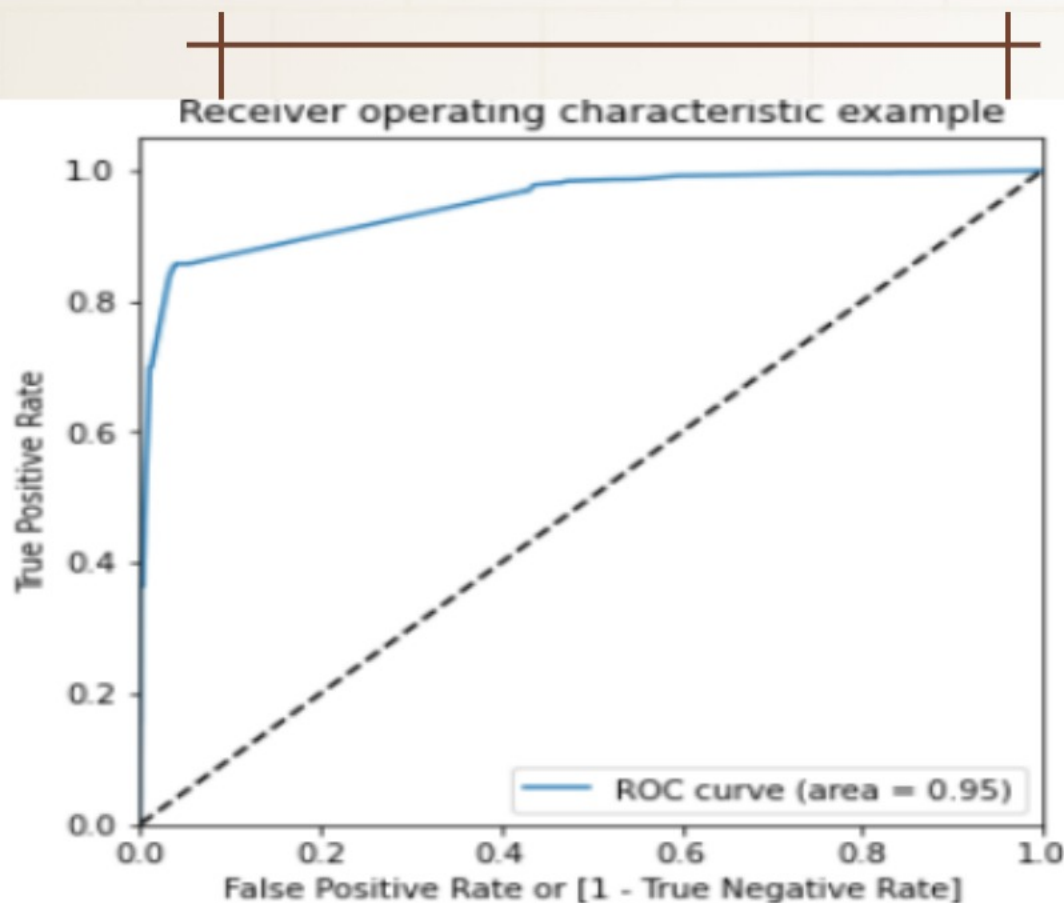


Last Activity by customers who have sms and email facility available have good conversion rates.



Specialization with the chart mentions converted and non converted data.

The count of others under specialization is more and can be filtered out.



Recommendations

Area under ROC is 0.95

Treating missing values and getting rid of outliers is important.

Remove the features which created imbalance during merging values.

Positive and false rates required for Gini of the Model.

GINI MODEL



“
Thank You
”
