



LEAD SCORING CASE STUDY

PRESENTED BY:
YESH THAKUR, EPGP DS C31



PROBLEM STATEMENT

- An Education Company named X Education sells online courses to Industry Professionals.
- Now, although X Education gets a lot of leads, it's lead conversion rate is very poor of around 30%.
- The company wants to increase it to 30%.



GOAL

- Build a Logistic Regression Model to assign a lead score between 0 to 100 to each of the leads which can be used by the company to get 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 most likely not get converted.



TOP FACTORS THAT IMPACT THE CONVERSION OF LEADS

Features	
7	Tags_Will revert after reading the email
1	Lead Origin_Lead Add Form
12	Lead Source_Welingak Website
9	Lead Source_Google
3	What is your current occupation_Working Profes...
6	Tags_Ringing
8	Lead Source_Direct Traffic
0	Total Time Spent on Website
10	Lead Source_Organic Search
5	Tags_Interested in other courses
4	Specialization_Travel and Tourism
2	What is your current occupation_Student
11	Lead Source_Referral Sites



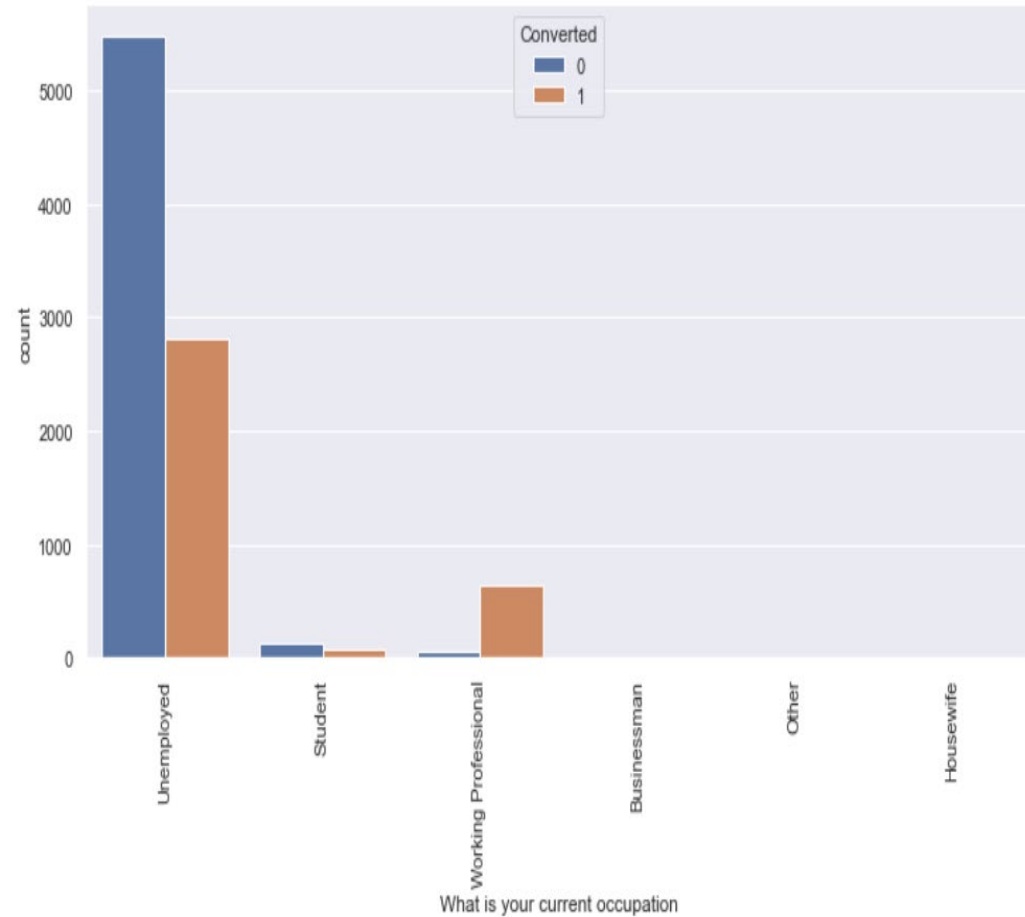
LET'S SEE
HOW THESE
FACTORS
ARE
IMPACTFUL

➡ WHAT MEASURES
SHOULD BE TAKEN
TO INCREASE THE
CONVERSION
RATE OF HOT
LEADS?

What is your current occupation

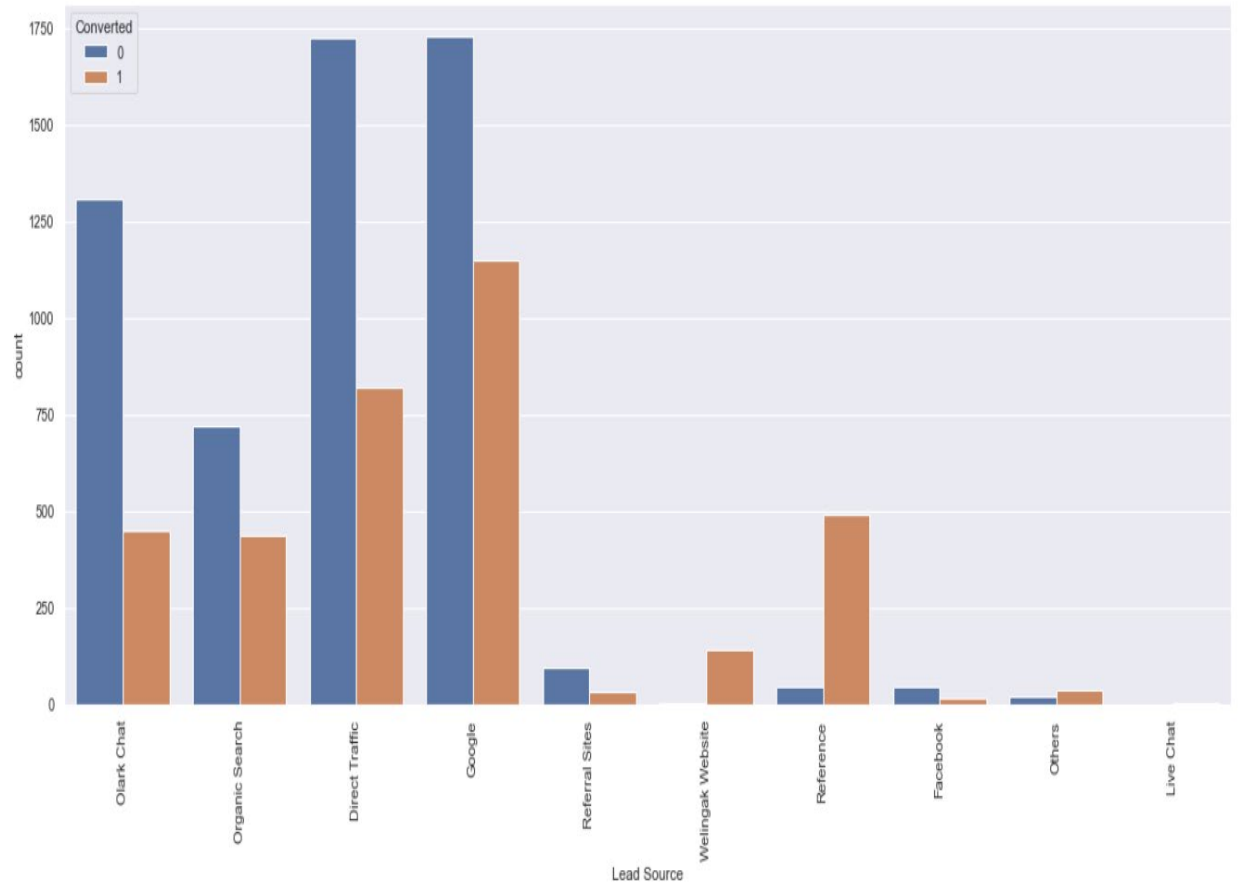
1. Unemployed leads are the maximum in the dataset which means they are most eager to learn something to get a job most probably.

2. Students and Working Professional leads have the higher chances of joining the course.



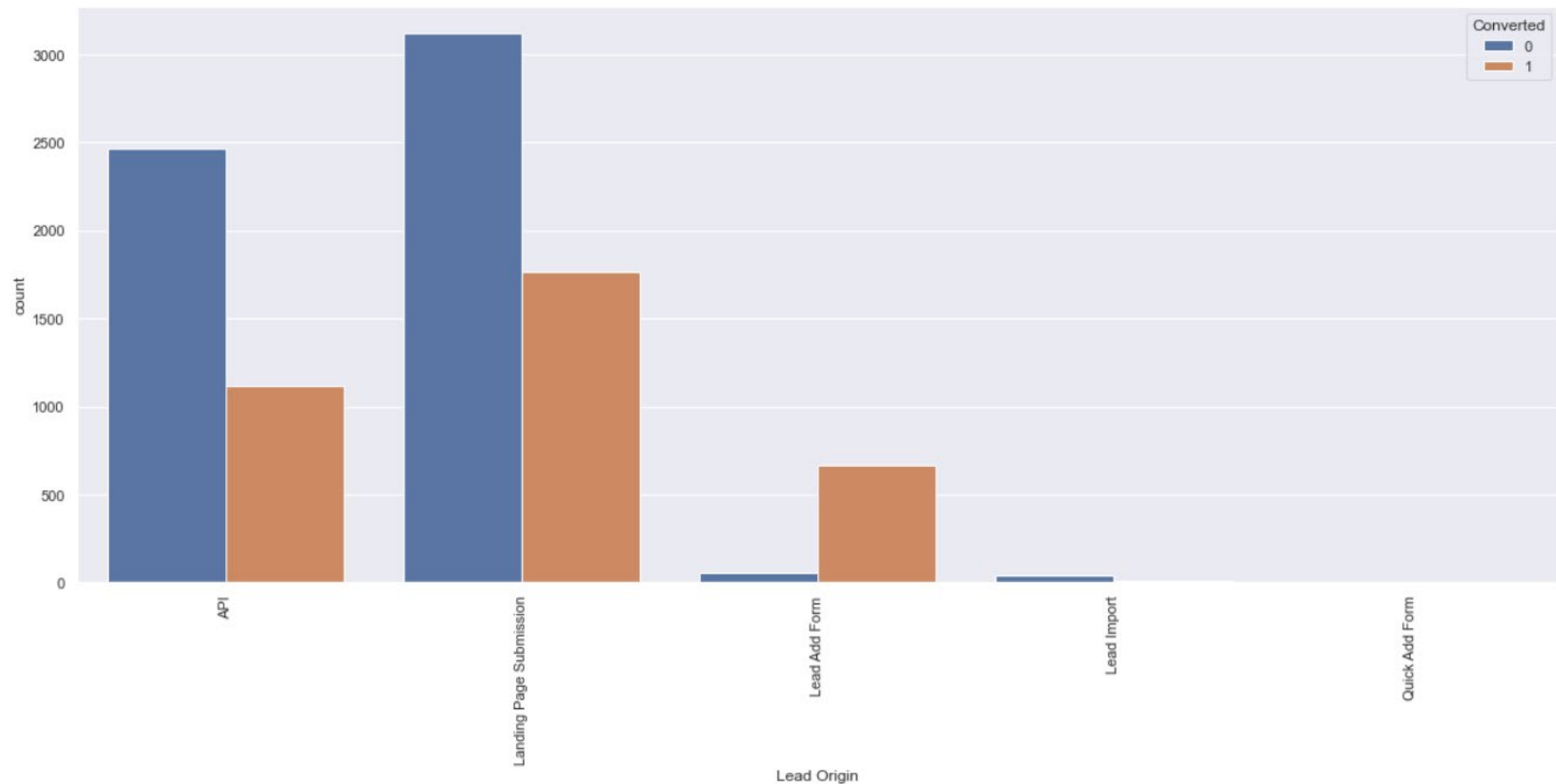
LEAD SOURCE

1. In order to increase the conversion, X Education should focus on generating more leads through Reference and Welingak Website. Also, focus on leads from Olark Chat, Organic Search, Direct Traffic and Google.
2. Conversion rate through Reference and Welingak Website is the highest.
3. Maximum leads are generated from Direct Traffic and Google.



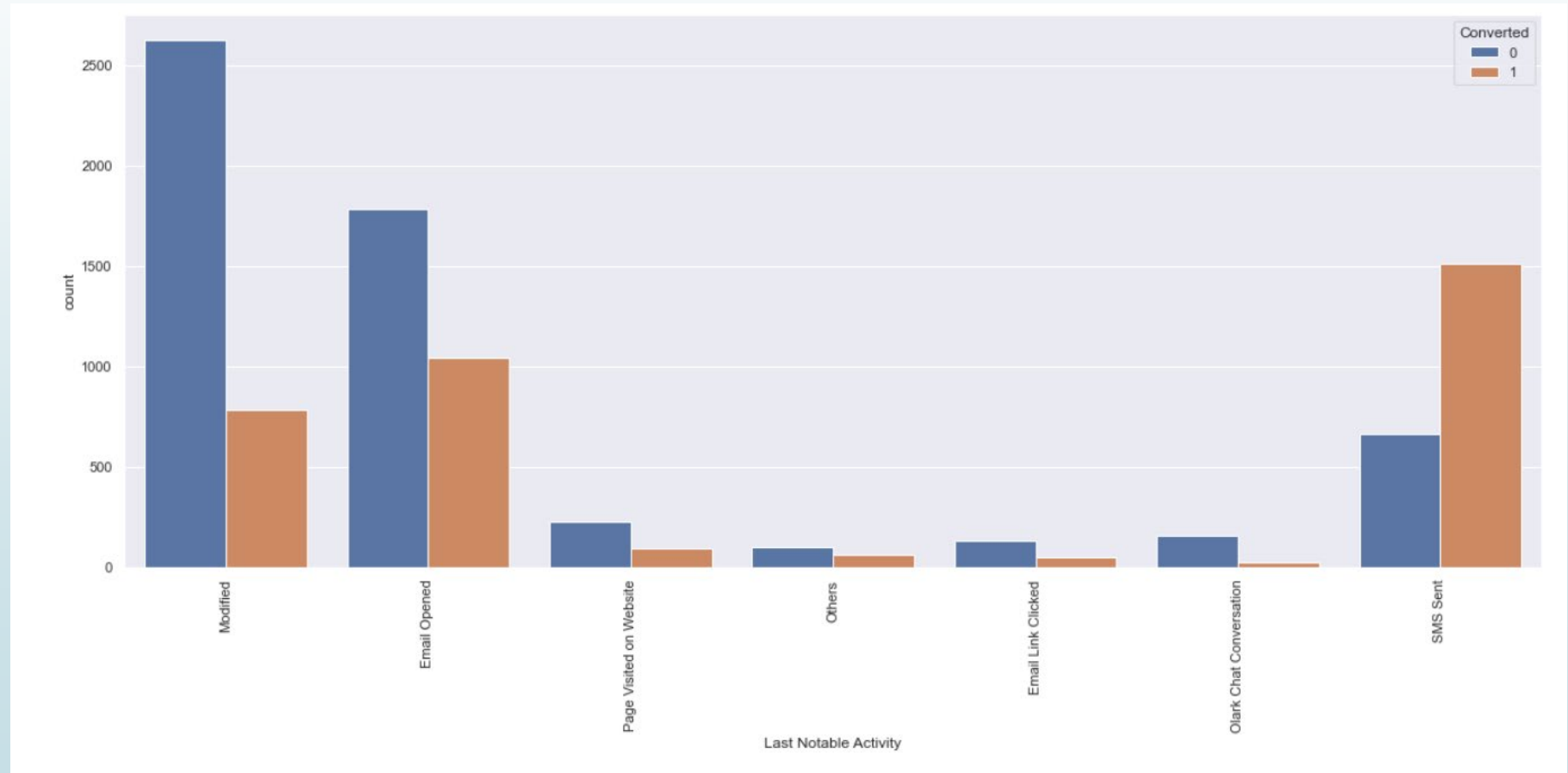
LEAD ORIGIN

In order to improve the conversion rate, X Education should focus on getting more leads through Lead Add Form and improve conversion of API and Landing Page Submission.



Last Notable Activity

Email Opened and Modified have the maximum leads, however the conversion is low so it should be focused on.

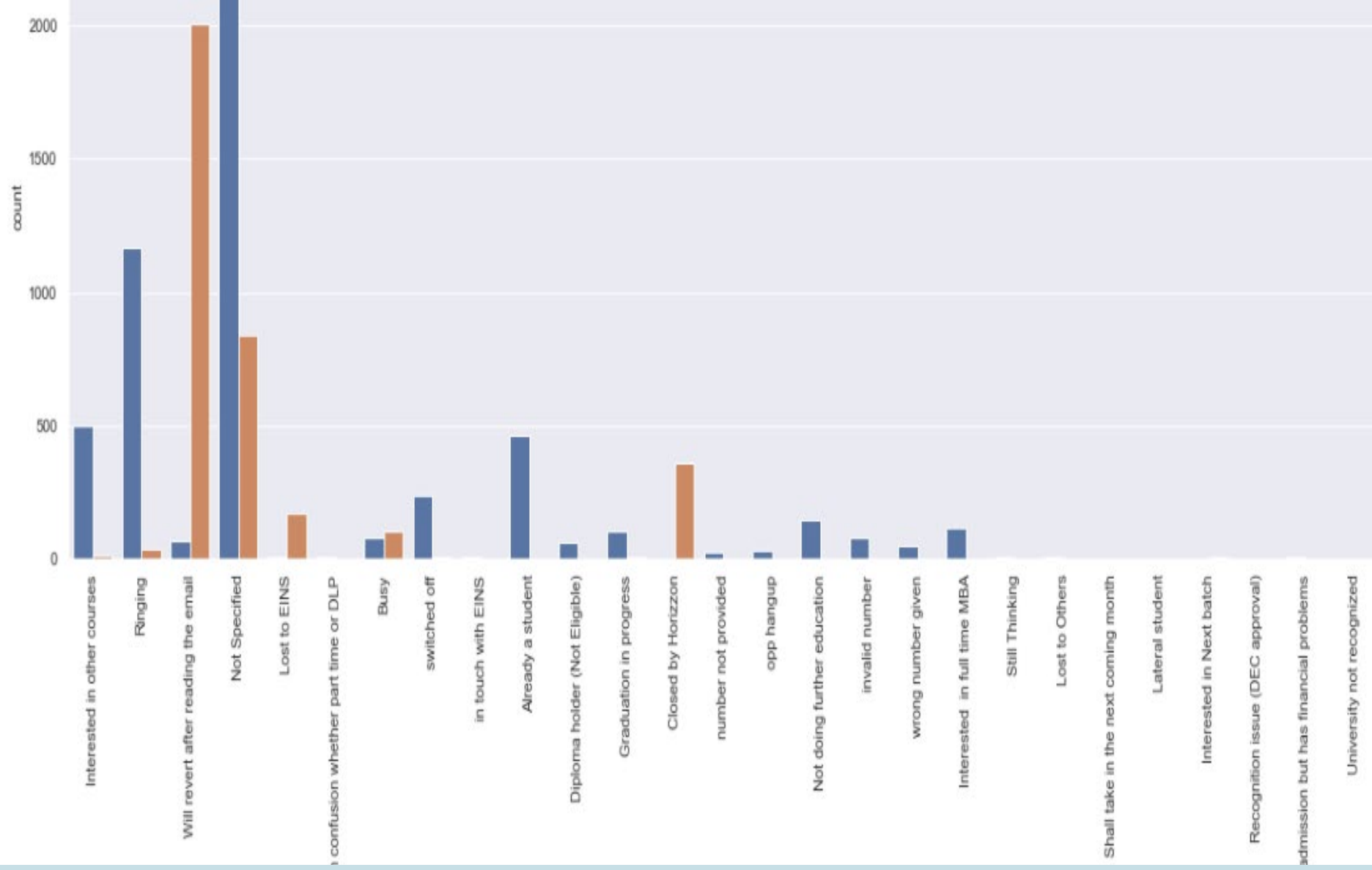




TAGS

- Lost to ENS
- Closed by Horizon

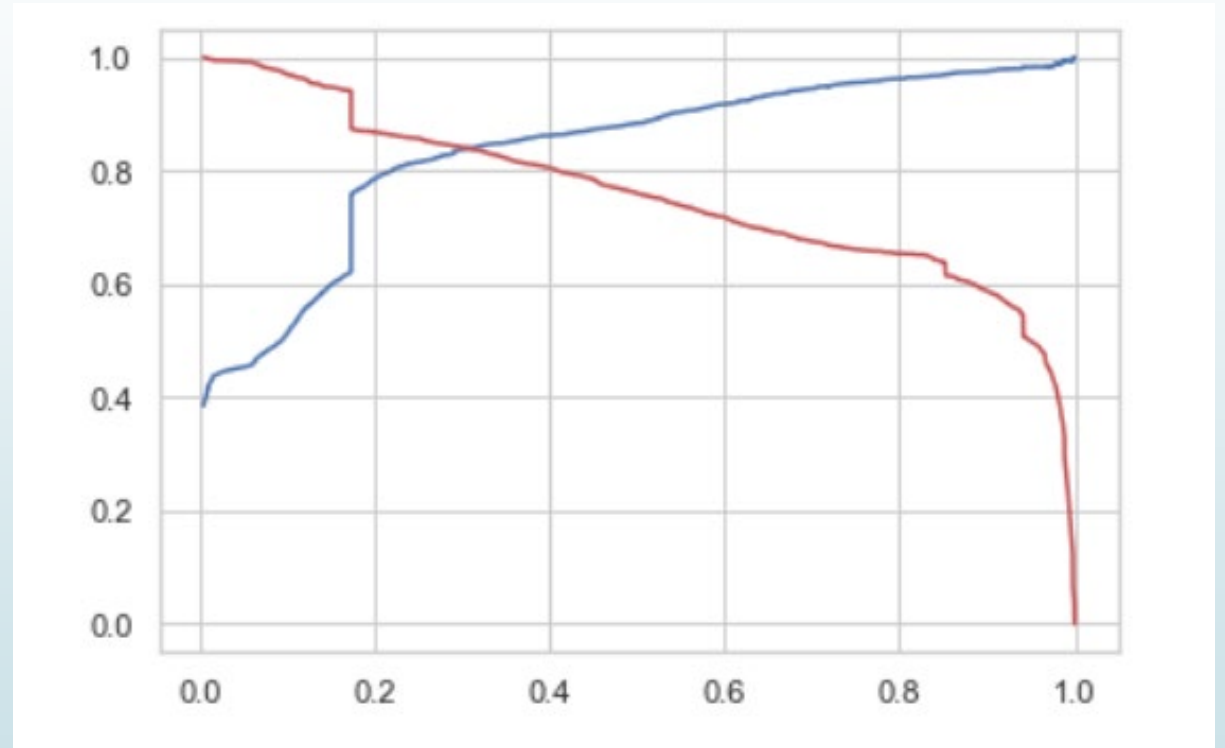
Have a very high conversion rate, highly correlated to Conversion.
Keeping an eye on these tags is important for conversion rate.



To achieve a conversion rate of around 80%

Need not to miss those prospects which can turn into hot leads which is Recall.

Should not over estimate a cold lead which is Precision.



Lead Prediction on the basis of probability set by the model

	Prospect ID	Converted	Converted_prob	Lead_Score	final_Predicted
0	7681	0	0.008303	1	0
1	984	0	0.121091	12	0
2	8135	0	0.281135	28	1
3	6915	0	0.009610	1	0
4	2712	1	0.927741	93	1

	Prospect ID	Converted	Converted_prob	Lead_Score	final_Predicted
2681	9029	1	0.173327	17	0
2682	1420	0	0.184568	18	0
2683	5996	1	0.104622	10	0
2684	4143	0	0.005907	1	0
2685	1005	1	0.292628	29	1



How 80% conversion is achieved?

- Even a lead that was not converted as per the available history, those leads could be guided properly based on the Lead Score provided, without wasting any resources on Low Probability leads.
- Sensitivity is the metric to show that the 80% conversion is achieved on the data.
- Any new prospect lead, generates a score of 38% in the model, then the company needs to take measure to guide the lead.