

- Which are the top three variables in your model which contribute most towards the probability of a lead getting converted?

Following three variables are contributing the most towards the probability of a lead getting converted:

- Lead Origin
- What is your current occupation
- Last Activity

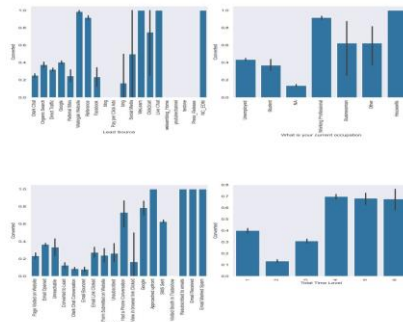
Reviewing coefficient of these variables indicate the importance of these variables.

	coef	std err	z	P> z	[0.025	0.975]
const	1.4827	0.195	7.596	0.000	1.100	1.865
Total Time Spent on Website	1.0521	0.044	23.748	0.000	0.965	1.139
Tags	0.0179	0.007	2.544	0.011	0.004	0.032
Lead Quality	-1.0473	0.045	-23.045	0.000	-1.136	-0.958
Asymmetrique Profile Index	0.1343	0.035	3.827	0.000	0.066	0.203
Asymmetrique Activity Score	0.7090	0.050	14.093	0.000	0.610	0.808
Asymmetrique Profile Score	0.1576	0.051	3.111	0.002	0.058	0.257
Lead Origin_Landing Page Submission	-0.5095	0.109	-4.657	0.000	-0.724	-0.295
Lead Origin_Lead Add Form	2.4765	0.209	11.825	0.000	2.066	2.887
Lead Source_Olark Chat	0.8050	0.135	5.966	0.000	0.541	1.069
Last Activity_Olark Chat Conversation	-1.6392	0.190	-8.609	0.000	-2.012	-1.266
Last Activity_SMS Sent	1.4649	0.085	17.219	0.000	1.298	1.632
What is your current occupation_NA	-0.9868	0.093	-10.595	0.000	-1.169	-0.804
What is your current occupation_Working Professional	1.6593	0.211	7.876	0.000	1.246	2.072

- What are the top 3 categorical/dummy variables in the model which should be focused the most on in order to increase the probability of lead conversion?

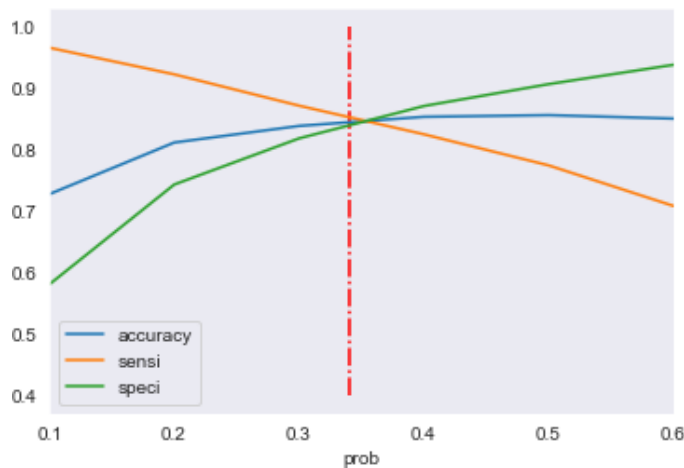
Upon reviewing the final model and as evident from EDA as well, followings are the top 3 categorical variables in the model which should be focused the most on in order to increase the probability of lead conversion:

- Lead Origin\_Lead Add Form – Lead has added the form.
- What is your current occupation\_Working Professional – Working professionals are more likely to take the course offered.
- Last Activity\_SMS Sent – SMS sent cases are more likely to convert



3. X Education has a period of 2 months every year during which they hire some interns. The sales team, in particular, has around 10 interns allotted to them. So during this phase, they wish to make the lead conversion more aggressive. So they want almost all of the potential leads (i.e. the customers who have been predicted as 1 by the model) to be converted and hence, want to make phone calls to as much of such people as possible. Suggest a good strategy they should employ at this stage.

Sensitivity with respect to our model can be defined as the ratio of total number of actual Conversions correctly predicted to the total number of actual conversions. Similarly, Specificity can be defined as the ratio of total number of actual non-conversions correctly predicted to the total number of actual non-conversions. For a model, as one increases, the other decreases and vice-versa. Different values of the sensitivity and specificity can be achieved for the same model by changing the conversion probability cutoff threshold value. For our model, the below graph shows how the Accuracy, Sensitivity and Specificity rating changes with change in the threshold value:



When the probability thresholds are very low, the sensitivity is very high and specificity is very low. Similarly, for larger probability thresholds, the sensitivity values are very low but the specificity values are very high.

High sensitivity implies that our model will correctly identify almost all leads who are likely to convert. It will do that by over-estimating the conversion likelihood, i.e. it will misclassify some non-conversion cases as conversions. Now, since X Education has more manpower for these 2 months and they wish to make the lead conversion more aggressive by wanting almost all the potential leads, we can choose a lower threshold value for Conversion Probability. This will ensure the Sensitivity rating is very high which in turn will make sure almost all leads who are likely to Convert are identified correctly and the agents can make phone calls to as much of such people as possible. The company may

follow high volume low margin strategy which means the conversion rate might reduce but, the count of conversion would increase and eventually the revenue would also increase.

4. Similarly, at times, the company reaches its target for a quarter before the deadline. During this time, the company wants the sales team to focus on some new work as well. So during this time, the company's aim is to not make phone calls unless it's extremely necessary, i.e. they want to minimize the rate of useless phone calls. Suggest a strategy they should employ at this stage.

Following the similar logic and context from the previous question, High Specificity implies that our model will correctly identify almost all leads who are not likely to convert. It will do that at the cost of losing out some low Conversion rate risky leads to the competition, i.e. it will misclassify some Conversion cases as non-Conversions.

Therefore, since X Education has already reached its target for a quarter and doesn't want to make phone calls unless it is extremely necessary, i.e. they want to minimize the rate of useless phone calls. We can choose a higher threshold value for Conversion Probability. This will ensure the Specificity rating is very high, which in turn will make sure almost all leads who are on the brink of the probability of getting Converted or not are not selected. As a result, they won't have to make unnecessary phone calls and can focus on some new work. In this way the efficiency of sales team would increase as the conversion rate would be high. The sales cycle will also be reduced.

