

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

absolute_coefficients	
What is your current occupation_Other	0.64
What is your current occupation_Working Professional	0.60
const	0.46
Lead Origin_Lead Add Form	0.39
Lead Source_Welingak Website	0.26
Last Notable Activity_Email Link Clicked	0.26
Last Notable Activity_Modified	0.24
What is your current occupation_Student	0.23
Last Notable Activity_Email Opened	0.22
Last Notable Activity_Page Visited on Website	0.22
Do Not Email	0.21
What is your current occupation_Unemployed	0.20
Last Activity_Olark Chat Conversation	0.18
Last Activity_Converted to Lead	0.14

Ans.

Figure 1 features that were used to build the model according to their importance (score)

So from the table given above, the top three variables in the model which contribute most towards the probability of a lead getting converted are:

- What is your current occupation_Other
- What is your current occupation_Working Professional
- Lead Origin_Lead Add Form

2. 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?

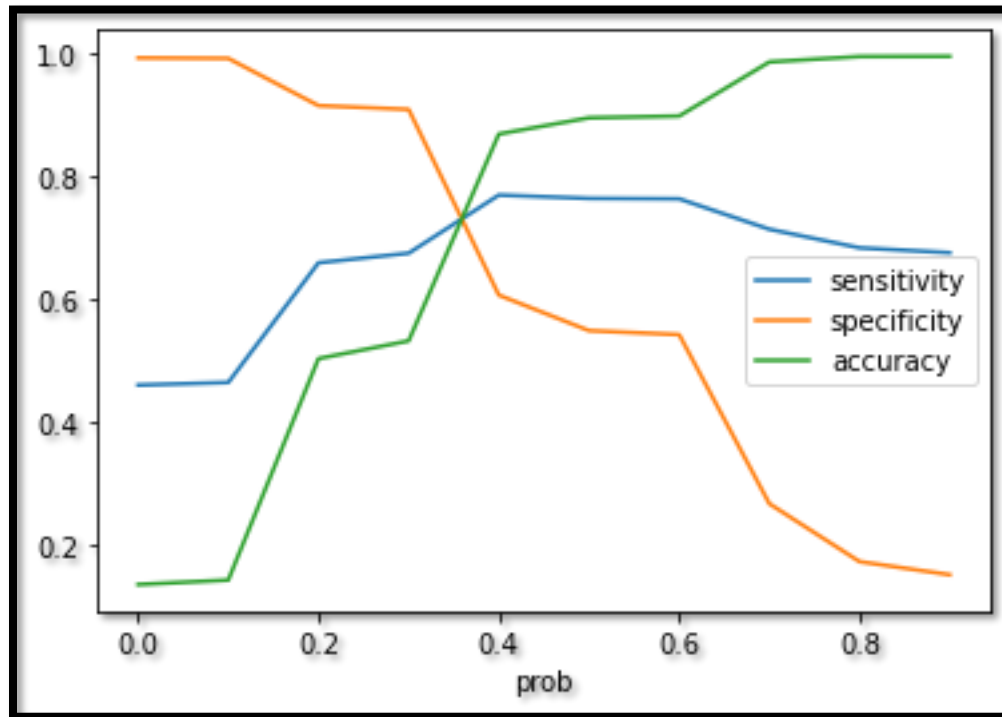
Ans. The answer is same as above. The top three categorical variables in the model which contribute most towards the probability of a lead getting converted are:

- What is your current occupation_Other
- What is your current occupation_Working Professional
- Lead Origin_Lead Add Form

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.

Ans. Looking at the conversion probabilities and our subsequent predictions, if we take a lower cutoff, lower than 0.37, more hot leads can be generated. And our model is more specific, and less sensitive. But given the trend the specificity and sensitivity follow, lower thresholds, will increase the sensitivity. This means that the non-conversion leads will be detected by the model.



So, it is a tradeoff between an increased conversion rate and lower sensitivity. But, given the fact that X-education has interns that can work on leads, this is feasible. So our suggestion would be to lower the threshold values.

- 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.**

Ans. According to our logical reasoning in the answer before, sensitivity is inversely related to specificity. This means that higher thresholds will have lower sensitivity, but high specificity. So higher specificity would mean that only actual hot leads are followed.

This will in turn increase the efficiency of the team, which is desired in this example. There will be no useless phone calls to leads that have a lower chance to be converted.