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

   Ans:

   The below image represents the features that are used to build the model in lead conversion as per their coefficient values.

	coef
const	-1.5092
Do Not Email	-1.0131
Time on Website	1.1066
Lead Origin_Landing Page Submission	-0.2612
Lead Origin_Lead Add Form	3.5675
Occupation_Working Professional	2.8604
Lead Source_Olark Chat	0.8593
Lead Source_Welingak Website	1.8485
Last Activity_Email Opened	0.5692
Last Activity_Less Popular Activity	2.0628
Last Activity_SMS Sent	1.7068
Last Notable Activity_Had a Phone Conversation	3.9275
Last Notable Activity_Modified	-0.8188
Last Notable Activity_Unreachable	2.0930
Specialization_Hospitality Management	-0.8389

   So, the top 3 variables that contribute most towards the probability of a lead getting converted are:

   1. Last Notable Activity_Had a Phone Conversation
   2. Lead Origin_Lead Add Form
   3. Occupation_Working Professional

2. What are the top 3 categorical/dummy variables in the model which get maximum focus in order to increase the probability of lead conversion?

   Ans:

   From the above diagram, the top 3 categorical/dummy variables that contribute the most towards the probability of a lead getting converted are also:

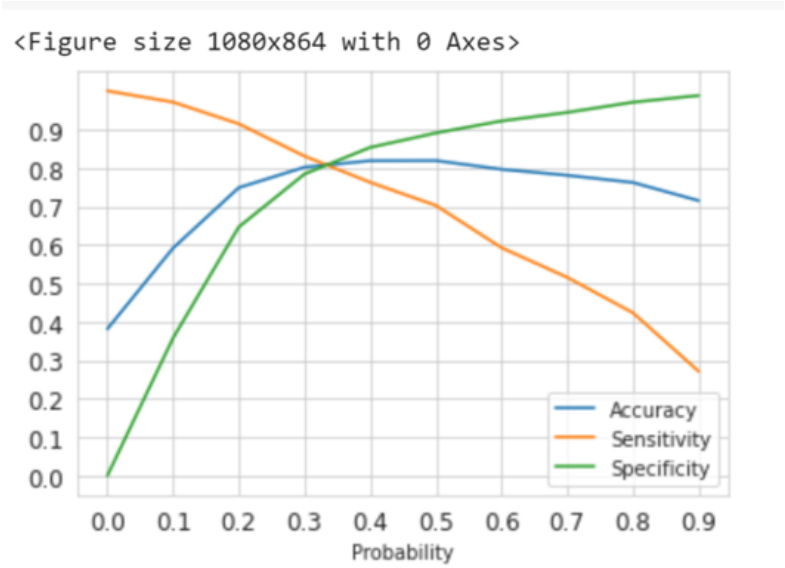
   1. Last Notable Activity_Had a Phone Conversation
   2. Lead Origin_Lead Add Form
   3. Occupation_Working Professional

3. X Education has a period of 2 months every year during which they hire few 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:

We are going to pick a lower threshold number for Conversion Probability. This will ensure that the Affectability rating is extremely high, ensuring that virtually all leads who are likely to switch over are recognized precisely and the operators can call as many of these people as possible.



When probability limits are very low, the sensitivity is high, and specificity is very low. Additionally, the sensitivity values are very low, but the specificity values are very high for greater probability thresholds.

High sensitivity implies that almost all leads who are likely to convert will be accurately identified by our model. Some non-Conversion instances will be incorrectly labelled as Conversions. i.e., overestimating the probability of conversion

We can choose a lower threshold value for Conversion Probability for these 2 months as X Education has abundant manpower and they wish to aggressively make the lead conversion by utilizing almost all of the potential leads,

This will make sure that the Sensitivity rating is extremely high which in turn ensures almost all leads who are likely to Convert are identified correctly and it will help the agents make the phone calls to as much of such people as possible.

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:

For Conversion Probability, a higher threshold value will be selected. This will guarantee a very high specificity rating, which in turn will guarantee that practically all leads who are on the verge of being selected for conversion or not are not. The agents can then concentrate on some fresh job since they won't have to make any pointless phone calls.

Converted	Converted_Probability	ID	Predict	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	Final_predicted_value	Lead_Score
0	0.257218	1871	0	1	1	1	0	0	0	0	0	0	0	0	26
0	0.232387	6795	0	1	1	1	0	0	0	0	0	0	0	0	23
0	0.300304	3516	0	1	1	1	1	0	0	0	0	0	0	0	30
0	0.809566	8105	1	1	1	1	1	1	1	1	1	1	0	1	81
0	0.132476	3934	0	1	1	0	0	0	0	0	0	0	0	0	13
1	0.992296	4844	1	1	1	1	1	1	1	1	1	1	1	1	99
0	0.118277	3297	0	1	1	0	0	0	0	0	0	0	0	0	12
1	0.978346	8071	1	1	1	1	1	1	1	1	1	1	1	1	98
0	0.160099	987	0	1	1	0	0	0	0	0	0	0	0	0	16
1	0.915495	7423	1	1	1	1	1	1	1	1	1	1	1	1	92
1	0.851301	1032	1	1	1	1	1	1	1	1	1	1	0	1	85
0	0.327542	6542	0	1	1	1	1	0	0	0	0	0	0	0	33
1	0.127663	4317	0	1	1	0	0	0	0	0	0	0	0	0	13
1	0.464057	6472	0	1	1	1	1	1	0	0	0	0	0	1	46
0	0.047338	712	0	1	0	0	0	0	0	0	0	0	0	0	5
0	0.257218	3960	0	1	1	1	0	0	0	0	0	0	0	0	26
0	0.257218	4654	0	1	1	1	0	0	0	0	0	0	0	0	26
0	0.015022	5902	0	1	0	0	0	0	0	0	0	0	0	0	2
0	0.257218	4691	0	1	1	1	0	0	0	0	0	0	0	0	26
1	0.202125	4341	0	1	1	1	0	0	0	0	0	0	0	0	20

High Specificity suggests that our model will properly identify nearly all leads who are not likely to convert, using a similar logic and context to the prior query. To achieve this, it will misclassify some conversion cases as non-conversions, hence losing out to the competition on some dangerous low conversion rate leads.

Consequently, since X Education has already surpassed its quarterly goal and wants to reduce the number of pointless phone calls, they can set a higher threshold value for Conversion Probability. This is because Education wants to avoid making phone calls until they are necessary.

This will guarantee a very high specificity rating, which in turn will guarantee that practically all leads who are on the verge of being selected or not for conversion are not. The agents can then concentrate on some fresh job since they won't have to make any pointless phone calls.