



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

➤ The top three variables in our model that contributes towards lead conversions are :-

1. Lead Origin_Lead Add Form
2. What is your current occupation_Working Professional
3. Total Time Spent on Website



	Features	VIF
2	Lead Origin_Lead Add Form	3.81
6	Lead Source_Reference	3.63
9	What is your current occupation_Unemployed	2.58
4	Lead Source_Google	1.70
3	Lead Source_Direct Traffic	1.67
5	Lead Source_Organic Search	1.31
10	What is your current occupation_Working Profes...	1.29
1	Total Time Spent on Website	1.12
8	What is your current occupation_Student	1.05
0	Do Not Email	1.03
7	Lead Source_Referral Sites	1.02



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?

➤ The top three variables that should be focused in our model are :-

1. Lead Origin_Lead Add Form
2. What is your current occupation_Working Professional
3. Total Time Spent on Website



	Features	VIF
2	Lead Origin_Lead Add Form	3.81
6	Lead Source_Reference	3.63
9	What is your current occupation_Unemployed	2.58
4	Lead Source_Google	1.70
3	Lead Source_Direct Traffic	1.67
5	Lead Source_Organic Search	1.31
10	What is your current occupation_Working Profes...	1.29
1	Total Time Spent on Website	1.12
8	What is your current occupation_Student	1.05
0	Do Not Email	1.03
7	Lead Source_Referral Sites	1.02



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.
- The final prediction is calculated based on a optimal cut off value of 0.3
In order to make the sales aggressive, the company may contact all the leads which have a conversion probability (value = 1) under a cut off 0.3

	Converted	Converted_prob	Prospect ID	Predicted	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	final_Predicted	Lead_Score
0	0	0.733427	7962	1	1	1	1	1	1	1	1	1	0	0	1	73
1	0	0.150019	5520	0	1	1	0	0	0	0	0	0	0	0	0	15
2	0	0.223565	1962	0	1	1	1	0	0	0	0	0	0	0	0	22
3	1	0.968245	1566	1	1	1	1	1	1	1	1	1	1	1	1	97
4	0	0.308725	9170	0	1	1	1	1	0	0	0	0	0	0	1	31
5	0	0.112311	5097	0	1	1	0	0	0	0	0	0	0	0	0	11
6	0	0.254701	8954	0	1	1	1	0	0	0	0	0	0	0	0	25
7	1	0.258253	309	0	1	1	1	0	0	0	0	0	0	0	0	26
8	1	0.112311	5519	0	1	1	0	0	0	0	0	0	0	0	0	11
9	1	0.648854	1050	1	1	1	1	1	1	1	1	0	0	0	1	65
10	0	0.095867	5246	0	1	0	0	0	0	0	0	0	0	0	0	10
11	0	0.065473	3280	0	1	0	0	0	0	0	0	0	0	0	0	7
12	0	0.151572	5758	0	1	1	0	0	0	0	0	0	0	0	0	15
13	0	0.134456	7389	0	1	1	0	0	0	0	0	0	0	0	0	13
14	1	0.308725	2912	0	1	1	1	1	0	0	0	0	0	0	1	31
15	1	0.057660	4204	0	1	0	0	0	0	0	0	0	0	0	0	6
16	0	0.234264	42	0	1	1	1	0	0	0	0	0	0	0	0	23
17	0	0.096431	250	0	1	0	0	0	0	0	0	0	0	0	0	10
18	1	0.064247	8534	0	1	0	0	0	0	0	0	0	0	0	0	6
19	0	0.023246	9161	0	1	0	0	0	0	0	0	0	0	0	0	2

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.

➤ A good strategy will be:

- To focus on narrow set of lead audience (discarding lower conversion probable leads)
- Technically, we can generate this new set of leads by altering (moving up) the value of cut off so as to discard lower conversion rate probable leads from our Logistic Regression Model
- Doing so, we will be doing minimal effort and still be getting fair conversions.