

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

Ans : As per Below screenshot, the top three variables that contribute most towards the probability of a lead getting converted based on coefficient values are :

- Lead Origin_Lead Add Form (with highest coef value of 3.3562)
- Current_Occupation_Working Professional
- Last Activity_SMS Sent

	coef	std err	z	P> z	[0.025	0.975]
const	-0.0323	0.090	-0.361	0.718	-0.208	0.143
Total Time Spent on Website	1.0906	0.040	27.528	0.000	1.013	1.168
Lead Origin_Lead Add Form	3.3562	0.239	14.028	0.000	2.887	3.825
Current_Occupation_Working Professional	2.7815	0.197	14.099	0.000	2.395	3.168
Lead Source_Direct Traffic	-1.4611	0.114	-12.864	0.000	-1.684	-1.238
Lead Source_Google	-1.0043	0.108	-9.284	0.000	-1.216	-0.792
Lead Source_Organic Search	-1.1360	0.132	-8.620	0.000	-1.394	-0.878
Lead Source_Referral Sites	-1.3556	0.337	-4.021	0.000	-2.016	-0.695
Last Activity_Email Bounced	-1.3929	0.318	-4.386	0.000	-2.015	-0.770
Last Activity_Olark Chat Conversation	-0.9400	0.195	-4.831	0.000	-1.321	-0.559
Last Activity_SMS Sent	1.1949	0.074	16.156	0.000	1.050	1.340
Last Notable Activity_Modified	-0.8610	0.083	-10.424	0.000	-1.023	-0.699
Last Notable Activity_Olark Chat Conversation	-1.1410	0.412	-2.768	0.006	-1.949	-0.333

- 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 : Based on the coefficient values from the above screenshot the following are the top three categorical/dummy variables which should be focused the most on in order to increase the probability of lead conversion are again same :

- Lead Origin_Lead Add Form
- Current_Occupation_Working Professional
- Last Activity_SMS Sent

- 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 : The final prediction is calculated based on an optimal cut off value of 0.33. In order to make the sales aggressive, the company may contact all the leads which have a conversion

	ID	Converted	Converted_Prob	Predicted	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	final_Predicted
0	9196	1	0.897882	1	1	1	1	1	1	1	1	1	0	1
1	4696	0	0.415869	0	1	1	1	1	0	0	0	0	0	1
2	3274	0	0.367702	0	1	1	1	0	0	0	0	0	0	1
3	2164	0	0.047136	0	0	0	0	0	0	0	0	0	0	0
4	1667	1	0.583175	1	1	1	1	1	1	0	0	0	0	1
...
6262	379	1	0.081448	0	0	0	0	0	0	0	0	0	0	0
6263	84	1	0.357334	0	1	1	1	0	0	0	0	0	0	1
6264	8313	0	0.095834	0	0	0	0	0	0	0	0	0	0	0
6265	7169	0	0.272181	0	1	1	0	0	0	0	0	0	0	0
6266	5825	1	0.998255	1	1	1	1	1	1	1	1	1	1	1

6267 rows x 14 columns

- Ans : In order to minimize the rate of useless phone calls, the company may contact all the leads which have a conversion probability (value = 1) (marked in yellow) under column 0.7 and 0.8. There can also be seen conflicts in some cases where leads are actually converted but model is showing non-convertible (marked in red). But these cannot have any impact as the target is already achieved.

	ID	Converted	Converted_Prob	Predicted	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	final_Predicted
0	9196	1	0.897882	1	1	1	1	1	1	1	1	1	0	1
1	4696	0	0.415869	0	1	1	1	1	0	0	0	0	0	1
2	3274	0	0.367702	0	1	1	1	0	0	0	0	0	0	1
3	2164	0	0.047136	0	0	0	0	0	0	0	0	0	0	0
4	1667	1	0.583175	1	1	1	1	1	1	0	0	0	0	1
...
6262	379	1	0.081448	0	0	0	0	0	0	0	0	0	0	0
6263	84	1	0.357334	0	1	1	1	0	0	0	0	0	0	1
6264	8313	0	0.095834	0	0	0	0	0	0	0	0	0	0	0
6265	7169	0	0.272181	0	1	1	0	0	0	0	0	0	0	0
6266	5825	1	0.998255	1	1	1	1	1	1	1	1	1	1	1

6267 rows x 14 columns