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

Solution:

As per the coefficient values from the below screenshot, the top three variables which contribute most towards the probability of a lead getting converted are:

- i) TotalVisits
- ii) Total Time Spent on Website
- iii) Lead Origin_Lead Add Form

	coef	std err	z	P> z	[0.025	0.975]
const	0.2040	0.196	1.043	0.297	-0.179	0.587
TotalVisits	11.1489	2.665	4.184	0.000	5.926	16.371
Total Time Spent on Website	4.4223	0.185	23.899	0.000	4.060	4.785
Lead Origin_Lead Add Form	4.2051	0.258	16.275	0.000	3.699	4.712
Lead Source_Olark Chat	1.4526	0.122	11.934	0.000	1.214	1.691
Lead Source_Welingak Website	2.1526	1.037	2.076	0.038	0.121	4.185
Do Not Email_Yes	-1.5037	0.193	-7.774	0.000	-1.883	-1.125
Last Activity_Had a Phone Conversation	2.7552	0.802	3.438	0.001	1.184	4.326
Last Activity_SMS Sent	1.1856	0.082	14.421	0.000	1.024	1.347
What is your current occupation_Student	-2.3578	0.281	-8.392	0.000	-2.908	-1.807
What is your current occupation_Unemployed	-2.5445	0.186	-13.699	0.000	-2.908	- 2.180
Last Notable Activity_Unreachable	2.7846	0.807	3.449	0.001	1.202	4.367

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?

Solution:

As per the coefficient values from the above screenshot, the top three categorical/dummy variables which contribute most towards the probability of a lead getting converted are:

- i) Lead Add Form (from Lead Origin)
- ii) Unreachable (from Last Notable Activity)
- iii) Had a Phone Conversation (from Last Activity)

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.

Solution:

So since the goal is to make the sales aggressive we will reduce the cutoff value so that more rows that are actually 1 will get predicted as 1 and not get wasted. We see by the image given below that 0.2 is the optimal reduction.

. ,	- 3 - 3													
	Converted	Conversion_Prob	Predicted	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	8.0	0.9	final_predicted
0	0	0.300117	0	1	1	1	1	0	0	0	0	0	0	0
1	0	0.142002	0	1	1	0	0	0	0	0	0	0	0	0
2	1	0.127629	0	1	1	0	0	0	0	0	0	0	0	0
3	1	0.291558	0	1	1	1	0	0	0	0	0	0	0	0
4	1	0.954795	1	1	1	1	1	1	1	1	1	1	1	1
5	0	0.194426	0	1	1	0	0	0	0	0	0	0	0	0
6	0	0.178073	0	1	1	0	0	0	0	0	0	0	0	0
7	1	0.949460	1	1	1	1	1	1	1	1	1	1	1	1
8	0	0.075995	0	1	0	0	0	0	0	0	0	0	0	0
9	1	0.982316	1	1	1	1	1	1	1	1	1	1	1	1
10	0	0.331556	0	1	1	1	1	0	0	0	0	0	0	0
11	0	0.117490	0	1	1	0	0	0	0	0	0	0	0	0
12	1	0.944915	1	1	1	1	1	1	1	1	1	1	1	1
13	1	0.448764	0	1	1	1	1	1	0	0	0	0	0	1
14	0	0.573896	1	1	1	1	1	1	1	0	0	0	0	1
15	1	0.987346	1	1	1	1	1	1	1	1	1	1	1	1

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.

Solution:

Since we want to reduce the rate of useless calls, we can increase the cutoff value to 0.8. That way the model predicts less calls as 1. But the problem will be that we may miss out on couple of useful calls too as shown in the image given below marked with red circles.

	Converted	Conversion_Prob	Predicted	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	8.0	0.9	final_predicted
0	0	0.300117	0	1	1	1	1	0	0	0	0	0	0	0
1	0	0.142002	0	1	1	0	0	0	0	0	0	0	0	0
2	1	0.127629	0	1	1	0	0	0	0	0	0	0	0	0
3	1	0.291558	0	1	1	1	0	0	0	0	0	0	0	0
4	1	0.954795	1	1	1	1	1	1	1	1	1	1	1	1
5	0	0.194426	0	1	1	0	0	0	0	0	0	0	0	0
6	0	0.178073	0	1	1	0	0	0	0	0	0	0	0	0
7	1	0.949460	1	1	1	1	1	1	1	1	1	1	1	1
8	0	0.075995	0	1	0	0	0	0	0	0	0	0	0	0
9	1	0.982316	1	1	1	1	1	1	1	1	1	1	1	1
10	0	0.331556	0	1	1	1	1	0	0	0	0	0	0	0
11	0	0.117490	0	1	1	0	0	0	0	0	0	0	0	0
12	1	0.944915	1	1	1	1	1	1	1	1	1	1	1	1
13	v	0.448764	0	1	1	1	1	1	0	0	0	0	0	1
14	0	0.573896	1	1	1	1	1	1	1	0	0	0	0	1
15	1	0.987346	1	1	1	1	1	1	1	1	1	1	1	1