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

Solution:

Based on the coefficient values from below screeshot, the following are the top three variables that contribute most towards the probability of a lead getting converted:

- a) Welingak Website (from Lead Source)
- b) Working Professional (from Current Occupation)
- c) Total Time Spent on Website

	coef
const	-2.6908
Do Not Email	-1.5288
TotalVisits	-0.5381
Total Time Spent on Website	3.7322
LeadOrigin_API	-0.0188
LeadOrigin_Lead Import	0.3242
LeadSource_Olark Chat	0.6801
LeadSource_Welingak Website	4.8295
LastActivity_Converted to Lead	-1.6025
LastActivity_Olark Chat Conversation	-1.3613
LastActivity_SMS Sent	1.2112
CurrentOccupation_Other	1.7281
CurrentOccupation_Student	1.5820
CurrentOccupation_Unemployed	1.3338
CurrentOccupation_Working Professional	4.0658
LastNotableActivity_Had a Phone Conversation	2.7735
LastNotableActivity_Unreachable	1.9341

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:

Again, based on the coefficient values from the screen shot in the question above, the following are the top three categorical/dummy variables that should be focused the most in order to increase the probability of lead conversion:

- a) Welingak Website (from Lead Source)
- b) Working Professional (from Current Occupation)
- c) Had a Phone Conversation (from Last Notable 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:

In the below image, the final prediction is calculated based on a optimal cut off value of 0.37.

In order to make the sales aggressive, the company may contact all the leads which have a conversion probabilty (value = 1) under a cut off 0.3 (column 0.3 highlighted in yellow).

	Converted	Converted_Prob	LeadId	predicted	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	8.0	0.9	final_predicted	lead_score
0	0	0.862679	7962	1	1	1	1	1	1	1	1	1	1	0	1	86
1	0	0.167977	5520	0	1	1	0	0	0	0	0	0	0	0	0	17
2	0	0.472414	1962	0	1	1	1	1	1	0	0	0	0	0	1	47
3	1	0.947484	1566	1	1	1	1	1	1	1	1	1	1	1	1	95
4	0	0.332745	9170	0	1	1	1	1	0	0	0	0	0	0	0	33
5	0	0.116135	5097	0	1	1	0	0	0	0	0	0	0	0	0	12
6	0	0.574341	8954	1	1	1	1	1	1	1	0	0	0	0	1	57
7	1	0.237591	309	0	1	1	1	0	0	0	0	0	0	0	0	24
8	1	0.306111	5519	0	1	1	1	1	0	0	0	0	0	0	0	31
9	1	0.586401	1050	1	1	1	1	1	1	1	0	0	0	0	1	59
10	0	0.203012	5246	0	1	1	1	0	0	0	0	0	0	0	0	20
11	0	0.064065	3280	0	1	0	0	0	0	0	0	0	0	0	0	6
12	0	0.206603	5758	0	1	1	1	0	0	0	0	0	0	0	0	21
13	0	0.146802	7389	0	1	1	0	0	0	0	0	0	0	0	0	15
14	1	0.626070	2912	1	1	1	1	1	1	1	1	0	0	0	1	63
15	1	0.050241	4204	0	1	0	0	0	0	0	0	0	0	0	0	5
16	0	0.215720	42	0	1	1	1	0	0	0	0	0	0	0	0	22
17	0	0.021001	250	0	1	0	0	0	0	0	0	0	0	0	0	2
18	1	0.067203	8534	0	1	0	0	0	0	0	0	0	0	0	0	7
19	0	0.022207	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.

Solution:

In order to minimize the rate of useless phone calls, the company may contact all the leads which have a conversion probabilty (value = 1 highlighted in yellow color) under column 0.7. However, the flipside here would be that, we may miss out on those leads that are actually converted but then the model wrongly predicted them as not converted. (See red highlights in the image below). This should not be a major cause for concern as the target has already be achieved.

	Converted	Converted_Prob	LeadId	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	0.862679	7962	1	1	1	1	1	1	1	1	1	1	0	1	86
1	0	0.167977	5520	0	1	1	0	0	0	0	0	0	0	0	0	17
2	0	0.472414	1962	0	1	1	1	1	1	0	0	0	0	0	1	47
3	1	0.947484	1566	1	1	1	1	1	1	1	1	1	1	1	1	95
4	0	0.332745	9170	0	1	1	1	1	0	0	0	0	0	0	0	33
5	0	0.116135	5097	0	1	1	0	0	0	0	0	0	0	0	0	12
6	0	0.574341	8954	1	1	1	1	1	1	1	0	0	0	0	1	57
7	R	0.237591	309	0	1	1	1	0	0	0	0	0)	0	0	0	24
8	[1]	0.306111	5519	0	1	1	1	1	0	0	0	0	0	0	0	31
9	1	0.586401	1050	1	1	1	1	1	1	1	0	0	0	0	1	59
10	0	0.203012	5246	0	1	1	1	0	0	0	0	0	0	0	0	20
11	0	0.064065	3280	0	1	0	0	0	0	0	0	0	0	0	0	6
12	0	0.206603	5758	0	1	1	1	0	0	0	0	0	0	0	0	21
13	0	0.146802	7389	0	1	1	0	0	0	0	0	0	0	0	0	15
14	1	0.626070	2912	1	1	1	1	1	1	1	1	0	0	0	1	63
15	1	0.050241	4204	0	1	0	0	0	0	0	0	0	0	0	0	5
16	0	0.215720	42	0	1	1	1	0	0	0	0	0	0	0	0	22
17	0	0.021001	250	0	1	0	0	0	0	0	0	0	0	0	0	2
18	1	0.067203	8534	0	1	0	0	0	0	0	0	0	0	0	0	7
19	0	0.022207	9161	0	1	0	0	0	0	0	0	0	0	0	0	2