# **Lead Scoring Assignment**

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

#### Answer: -

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

- a) Total Time Spent on Website
- b) Lead Add Form (from Lead Origin)
- c) Had a Phone Conversation (from Last Notable Activity)
- 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?

## Answer: -

Again, based on the coefficient values, 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) Lead Add Form (from Lead Origin)
- b) Had a Phone Conversation (from Last Notable Activity)
- c) Working Professional (from What is your current occupation)

coef	
-2.3134	const
-1.1067	Do Not Email
1.5038	TotalVisits
4.5275	Total Time Spent on Website
-0.3253	LeadOrigin_Landing Page Submission
3.7313	LeadOrigin_Lead Add Form
1.5706	LeadSource_Olark Chat
2.1406	LeadSource_Welingak Website
-0.9859	LastActivity_Email Bounced
-1.5943	LastActivity_Not Sure
-1.3427	LastActivity_Olark Chat Conversation
1.3544	LastActivity_SMS Sent
-1.3041	CurrentOccupation_No Information
2.6421	CurrentOccupation_Working Professional
3.2904	LastNotableActivity_Had a Phone Conversation
1.7567	LastNotableActivity_Unreachable

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.

#### Answer: -

The final prediction is calculated based on the optimal cut off value of **0.37** in the image below.

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	0.600894	2240	1	1	1	1	1	1	1	1	0	0	0	1	60
1	0	0.794876	113	1	1	1	1	1	1	1	1	1	0	0	1	79
2	1	0.802664	4132	1	1	1	1	1	1	1	1	1	1	0	1	80
3	0	0.089476	5573	0	1	0	0	0	0	0	0	0	0	0	0	9
4	0	0.029809	1109	0	1	0	0	0	0	0	0	0	0	0	0	3
5	0	0.041753	2282	0	1	0	0	0	0	0	0	0	0	0	0	4
6	1	0.767093	2976	1	1	1	1	1	1	1	1	1	0	0	1	77
7	0	0.418665	8431	0	1	1	1	1	1	0	0	0	0	0	1	42
8	1	0.805010	2770	1	1	1	1	1	1	1	1	1	1	0	1	81
9	1	0.995568	5790	1	1	1	1	1	1	1	1	1	1	1	1	100
10	1	0.950498	2943	1	1	1	1	1	1	1	1	1	1	1	1	95
11	0	0.197985	1196	0	1	1	0	0	0	0	0	0	0	0	0	20
12	1	0.648304	8874	1	1	1	1	1	1	1	1	0	0	0	1	65
13	0	0.106911	1491	0	1	1	0	0	0	0	0	0	0	0	0	11
14	0	0.110507	7676	0	1	1	0	0	0	0	0	0	0	0	0	11
15	1	0.490611	8750	0	1	1	1	1	1	0	0	0	0	0	1	49
16	1	0.805010	5049	1	1	1	1	1	1	1	1	1	1	0	1	81
17	0	0.648304	5691	1	1	1	1	1	1	1	1	0	0	0	1	65
18	1	0.380602	5773	0	1	1	1	1	0	0	0	0	0	0	1	38
19	0	0.032619	3906	0	1	0	0	0	0	0	0	0	0	0	0	3

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.

## Answer: -

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	8.0	0.9	final_predicted	lead_score
0	0	0.600894	2240	1	1	1	1	1	1	1	1	0	0	0	1	60
1	0	0.794876	113	1	1	1	1	1	1	1	1	1	0	0	1	79
2	1	0.802664	4132	1	1	1	1	1	1	1	1	1	1	0	1	80
3	0	0.089476	5573	0	1	0	0	0	0	0	0	0	0	0	0	9
4	0	0.029809	1109	0	1	0	0	0	0	0	0	0	0	0	0	3
5	0	0.041753	2282	0	1	0	0	0	0	0	0	0	0	0	0	4
6	1	0.767093	2976	1	1	1	1	1	1	1	1	1	0	0	1	77
7	0	0.418665	8431	0	1	1	1	1	1	0	0	0	0	0	1	42
8	1	0.805010	2770	1	1	1	1	1	1	1	1	1	1	0	1	81
9	1	0.995568	5790	1	1	1	1	1	1	1	1	1	1	1	1	100
10	1	0.950498	2943	1	1	1	1	1	1	1	1	1	1	1	1	95
11	0	0.197985	1196	0	1	1	0	0	0	0	0	0	0	0	0	20
12	1	0.648304	8874	1	1	1	1	1	1	1	1	0	0	0	1	65
13	0	0.106911	1491	0	1	1	0	0	0	0	0	0	0	0	0	11
14	0	0.110507	7676	0	1	1	0	0	0	0	0	0	0	0	0	11
15	1	0.490611	8750	0	1	1	1	1	1	0	0	0	<b>7</b> 0	0	1	49
16	1	0.805010	5049	1	1	1	1	1	1	1	1	1	1	0	1	81
17	0	0.648304	5691	1	1	1	1	1	1	1	1	0	0	0	1	65
18	1	0.380602	5773	0	1	1	1	1	0	0	0	0	0	0	1	38
19	0	0.032619	3906	0	1	0	0	0	0	0	0	0	0	0	0	3