

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 screenshot, the following are the top three variables that contribute most towards the probability of a lead getting converted :

- a) Welingak Website
- b) Reference
- c) Working Professional (from What is your current occupation)

	coef
const	1.3380
Do Not Email	-1.6086
Total Time Spent on Website	1.0959
Lead Origin_Landing Page Submission	-1.1641
Lead Source_Olark Chat	1.0618
Lead Source_Reference	3.3261
Lead Source_Welingak Website	5.9126
Last Activity_Converted to Lead	-1.0641
Last Activity_Email Bounced	-1.2351
Last Activity_Olark Chat Conversation	-1.3474
Specialization_Unknown	-1.1499
What is your current occupation_Working Professional	2.6124
Last Notable Activity_Email Link Clicked	-1.8946
Last Notable Activity_Email Opened	-1.4457
Last Notable Activity_Modified	-1.7558
Last Notable Activity_Olark Chat Conversation	-1.4783
Last Notable Activity_Page Visited on Website	-1.7218

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) Reference
- b) Working Professional (from What is your current occupation)
- c) Olark Chat

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 0.34.

In order to make the sales aggressive, the company may contact all the leads which have a conversion probabiltly (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	0.190392	3009	0	1	1	0	0	0	0	0	0	0	0	0	19
1	0	0.141953	1012	0	1	1	0	0	0	0	0	0	0	0	0	14
2	0	0.185466	9226	0	1	1	0	0	0	0	0	0	0	0	0	19
3	1	0.878116	4750	1	1	1	1	1	1	1	1	1	1	0	1	88
4	1	0.652278	7987	1	1	1	1	1	1	1	1	0	0	0	1	65

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 probability (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. This should not be a major cause for concern as the target has already been achieved.

	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	0.190392	3009	0	1	1	0	0	0	0	0	0	0	0	0	19
1	0	0.141953	1012	0	1	1	0	0	0	0	0	0	0	0	0	14
2	0	0.185466	9226	0	1	1	0	0	0	0	0	0	0	0	0	19
3	1	0.878116	4750	1	1	1	1	1	1	1	1	1	1	1	1	88
4	1	0.652278	7987	1	1	1	1	1	1	1	1	0	0	0	1	65

Submitted by DSC 65:

Rafida Khanum

Keshavi Gupta

Venkat Aditya