

- 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 :

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

Dep. Variable:	Converted	No. Observations:	6320
Model:	GLM	Df Residuals:	6302
Model Family:	Binomial	Df Model:	17
Link Function:	Logit	Scale:	1.0000
Method:	IRLS	Log-Likelihood:	-2514.1
Date:	Tue, 16 Jan 2024	Deviance:	5028.2
Time:	21:13:46	Pearson chi2:	6.30e+03
No. Iterations:	7	Pseudo R-squ. (CS):	0.4181
Covariance Type:	nonrobust		

	coef	std err	z	P> z	[0.025	0.975]
const	-2.5174	0.152	-16.541	0.000	-2.816	-2.219
Do Not Email	-1.1242	0.175	-6.414	0.000	-1.468	-0.781
TotalVisits	1.7711	0.289	6.122	0.000	1.204	2.338
Total Time Spent on Website	4.5135	0.169	26.784	0.000	4.183	4.844
Page Views Per Visit	-0.8513	0.250	-3.409	0.001	-1.341	-0.362
Search	-1.6344	1.011	-1.617	0.106	-3.615	0.346
LeadOrigin_Lead Add Form	3.0307	0.511	5.931	0.000	2.029	4.032
LeadSource_Olark Chat	1.4756	0.134	11.022	0.000	1.213	1.738
LeadSource_Reference	0.6636	0.539	1.231	0.219	-0.393	1.721
LeadSource_Welingak Website	2.6507	0.875	3.029	0.002	0.935	4.366
LastActivity_Email Opened	0.5072	0.116	4.373	0.000	0.280	0.735
LastActivity_Olark Chat Conversation	-0.6397	0.189	-3.383	0.001	-1.010	-0.269
LastActivity_SMS Sent	1.7015	0.117	14.517	0.000	1.472	1.931
CurrentOccupation_No Information	-1.2474	0.090	-13.845	0.000	-1.424	-1.071
CurrentOccupation_Working Professional	2.6117	0.203	12.860	0.000	2.214	3.010
LastNotableActivity_Had a Phone Conversation	3.5194	1.177	2.989	0.003	1.212	5.827
LastNotableActivity_Modified	-0.5343	0.090	-5.961	0.000	-0.710	-0.359
LastNotableActivity_Unreachable	2.0066	0.556	3.606	0.000	0.916	3.097

- 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 :

- Lead Add Form (from Lead Origin)
- Had a Phone Conversation (from Last Notable Activity)
- Working Professional (from What is your current occupation)

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.

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. (See red highlights in the image below). This should not be a major cause for concern as the target has already be achieved.