

Subjective Questions:

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

ANS:

	coef	std err	z	P> z	[0.025	0.975]
const	-1.4240	0.133	-10.703	0.000	-1.685	-1.163
Do Not Email	-1.7241	0.187	-9.238	0.000	-2.090	-1.358
Total Time Spent on Website	4.5093	0.165	27.370	0.000	4.186	4.832
Lead Origin_Landing Page Submission	-1.1073	0.127	-8.729	0.000	-1.356	-0.859
Lead Origin_Lead Import	1.1573	0.479	2.418	0.016	0.219	2.095
Lead Source_Olark Chat	0.9344	0.119	7.877	0.000	0.702	1.167
Lead Source_Reference	3.5340	0.243	14.518	0.000	3.057	4.011
Lead Source_Welingak Website	6.1637	0.734	8.397	0.000	4.725	7.602
Last Notable Activity_Had a Phone Conversation	3.5502	1.099	3.230	0.001	1.396	5.704
Last Notable Activity_SMS Sent	1.7876	0.079	22.523	0.000	1.632	1.943
Last Notable Activity_Unreachable	1.9022	0.474	4.014	0.000	0.973	2.831
Last Notable Activity_Unsubscribed	1.8227	0.526	3.463	0.001	0.791	2.854
Specialization_others	-1.2517	0.123	-10.178	0.000	-1.493	-1.011
What is your current occupation_Working Professional	2.6250	0.194	13.506	0.000	2.244	3.006

From the above model we can make out the top 3 variables contributing towards the probability of a lead getting converted are:

- Lead Source_Welingak Website
- Total time spent on website
- Last Notable Activity_Had a Phone Conversation

2. What are the top 3 categorical/dummy variables in the model which get maximum focus in order to increase the probability of lead conversion?

ANS:

The top 3 categorical/dummy variables that should be focused on to increase the probability of the lead conversion are:

- Lead Source_Welingak Website
- Last Notable Activity_Had a Phone Conversation
- Lead Source_Reference

The 3 other parameter which should be decreased in order to increase the conversion rate are:

- Lead Origin_Landing Page Submission
- Specialization – Others (those who didn't select any Specialization)
- Do not Email

3. X Education has a period of 2 months every year during which they hire few 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.

ANS:

In our scenario, we need to increase the chances of correctly predicting the lead conversions (1's) – which means we need to increase the Sensitivity of the model.

So, in order to make sales more aggressive, **we can lower the cut-off value** (currently set at 0.35 in our model) so that the model predicts more of the leads as "hot leads" – so the interns would make more sales calls and the lead conversion can be more aggressive.

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.

ANS:

In order to minimize the rate of the useless phone calls we can focus on reducing the False Positive Rate so that we can reduce the number of leads being classified as "hot leads" which are actually not very likely to convert.

In other words, we can focus on bringing our Specificity higher and thus, reducing the misclassifications. This **can be achieved if we increase our current threshold** (0.35 as per our final model) so that only the very high probabilistic leads are targeted and we reduce on the unnecessary phone calls.