Lead Score Assignment Subjective Questions and Answers

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

Answer-1 : As given below Model screen shot , we can see the top 3 variables, where coefficient are relatively high than others variable contributing most towards the probability of a lead gettiing converted.

Top 3 Variables
1- Lead Source_Welingak Website
2- Lead Source_Reference
3- Last Notable Activity_Had a Phone Conversation

	coef	std err	z	P>|z|	[0.025	0.975]
const	1.5546	0.182	8.541	0.000	1.198	1.911
Do Not Email	-1.2251	0.165	-7.442	0.000	-1.548	-0.902
Total Time Spent on Website	1.0782	0.039	27.723	0.000	1.002	1.154
Lead Source_Olark Chat	1.1110	0.101	10.987	0.000	0.913	1.309
Lead Source_Others	1.2840	0.532	2.414	0.016	0.241	2.327
Lead Source_Reference	3.7841	0.205	18.448	0.000	3.382	4.186
Lead Source_Welingak Website	5.5969	0.723	7.740	0.000	4.180	7.014
Last Activity_Olark Chat Conversation	-0.6938	0.194	-3.574	0.000	-1.074	-0.313
Last Activity_Others	0.8872	0.251	3.541	0.000	0.396	1.378
Last Activity_SMS Sent	1.2765	0.074	17.343	0.000	1.132	1.421
Specialization_Hospitality Management	-0.9073	0.321	-2.830	0.005	-1.536	-0.279
What is your current occupation_Other	-2.4908	0.810	-3.076	0.002	-4.078	-0.904
What is your current occupation_Student	-2.3648	0.292	-8.100	0.000	-2.937	-1.793
What is your current occupation_Unemployed	-2.7703	0.182	-15.224	0.000	-3.127	-2.414
Last Notable Activity_Had a Phone Conversation	2.5380	1.125	2.257	0.024	0.334	4.742
Last Notable Activity_Modified	-0.9267	0.082	-11.281	0.000	-1.088	-0.766
Last Notable Activity_Olark Chat Conversation	-0.6490	0.368	-1.762	0.078	-1.371	0.073

Q.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 - 2 : The top 3 Categorical /dummy Variable are given below which should be focused the most on in order to increase the probability of lead conversion.

1. Lead Source_Welingak Website
2. Lead Source_Reference
1. Last Notable Activity_Had a Phone Conversation

Q.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-3 : We can handle the situation using sensitivity concept also known as True positive rate. Sensitivity measure how well the Model predict the positive instances , so we require High sensitivity to predict correctly all the leads who are likely to convert.

We can increase the sensitivity by changing the model cutoff threshold for probability of lead conversion as given below.We should note that with increase in sensitivity the model may predict some of misclasify non conversion as conversion which company can cop up easily with their hired sales interns during the period.

So company need to use low cutoff threshold to get high sensitivity and making lead conversion more aggressive and can make phone calls to all the potential lead (value = 1)as much as possible.

As we can observe from below table company can use low threshold cutoff to get high sensitivity.

	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	0.256044	1871	0	1	1	1	0	0	0	0	0	0	0	0	26
1	0	0.228377	6795	0	1	1	1	0	0	0	0	0	0	0	0	23
2	0	0.297715	3516	0	1	1	1	0	0	0	0	0	0	0	0	30
3	0	0.820736	8105	1	1	1	1	1	1	1	1	1	1	0	1	82

4   0   0. 123677   3934   0   1   1   0   0   0   0   0   0   0   0   0   12

Q.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-4 : This is just opposite situation from above.Under such situation we need to use Specificity concept , which is also know as True Negative Rate. Specificity measures model ability to predict the true negative instances. Higher the specificity of the model indicates how correclty model identify most of negative results.So we can achieve high spacificity by increasing the cutoff threshold for probability of lead conversion.

Sametime we also need to note that with high specificity , model may predict some of conversion misclasify as non conversion, but since the company already reached its target , and wanting to focus only on those leads where probabiltity of conversion is high. So company should use high Specificity with high cutoff threshold of probability conversion to ensure that sales team not making any call unless it is extremely necessary.

From below plot we can see the change in cutoff and change in Accuracy , Specificity and Sensitivity.