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

Answer:

If we see the coefficients of all the features of final model, we would find that –

Features	Coeff
Tags_Will revert after reading the email	4.3
Lst_Ntbl_Act_SMS Sent	1.83
Total Time Spent on Website	0.99
Lst_Act_L_A_Otr	-0.63
Ld_Sr_Olark Chat	-2.18
Ld_Sr_Google	-2.79
Ld_Sr_Organic Search	-2.96
Ld_Sr_Direct Traffic	-3.07
Tags_Ringing	-3.31

Now considering top-3 variables (keep in mind that the contribution can be either positive or negative) –

Features	Coeff
Tags_Will revert after reading the email	4.3
Tags_Ringing	-3.31
Ld_Sr_Direct Traffic	-3.07

How to interpret Positive Coefficients?

A positive coefficient simply implies that the probability that the event identified by the Dependent Variable happens increases as the value of the Independent Variable increases. In other words, when the value of the Independent Variable increases the probability increases.

How to interpret Negative Coefficients?

A negative coefficient simply implies that the probability that the event identified by the Dependent Variable happens decreases as the value of the Independent Variable increases. In other words, when the value of the Independent Variable increases the probability decreases.

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 the contribution to probability can be positive or negative. If we minimize the value of independent variable to increase the overall Probability of Dependent variable.

   The answer remains same as previous one. All the 3 variables which contribute most to the model are Dummy Variables.

   So, when the Tag is as **"Tags_Will revert after reading the email"** then there is a high chance of Lead to get converted.
   However, when the Tag is **"Tags_Ringing"** its quite opposite and hence Ringing Tags should be focused as those are not good Leads.
   Same applies with **"Ld_Sr_Direct Traffic"** dummy variable. Such Leads should be focused and not included to target for conversion.

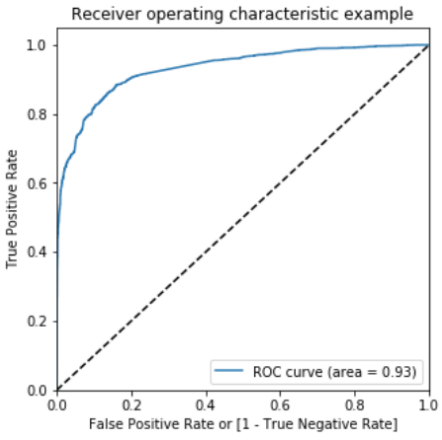
3. X Education has a period of 2 months every year during which they hire some interns. The sales team, in particular, has around 10 interns allotted to them. So during this phase, they wish to make the lead conversion more aggressive. So they want almost all of the potential leads (i.e. the customers 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:

   Here in the given case we should care about TPR (Sensitivity) more because we have extra resources to put and we really care don't care much about accuracy.

   Now if we go with lower cut-off (say 10%), my sensitivity reaches to almost 96% and it means we are hitting almost all the potential Leads.

Probability	FPR (1-Specificity)	TPR(Sensitivity)
0.00	1.00	1.00
0.10	0.50	0.96
0.20	0.18	0.89
0.30	0.14	0.87
0.40	0.10	0.82
0.50	0.07	0.75
0.60	0.04	0.69
0.70	0.03	0.66
0.80	0.02	0.63
0.90	0.01	0.54
1.00	0.00	0.00



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:

   Please refer above FPR/TPR matrix and ROC.

   To minimize the rate of useless phone call, company should focus more on minimizing FPR. It means the probability Cut-Off should be on higher end. However, it will make many potential leads uncontacted (which is agreed by business as given in above problem statement). If we take probability Cut-Off as 90% then we still able to target around 54% positive leads having less than 1% chance that the lead eventually does not convert. This may be the best situation to convert almost every Lead as a Customer.