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

ANS.

TOP 3 Features Contributing Most towards Probability
1. Lead Origin.
2. What is your current occupation.
3. Last Notable Activity

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?**

ANS

TOP 3 Categorical_Dummy variables in the model, contributing most
1. Lead Origin_Lead Add Form
2. What is your current occupation_Working Professional
3. Last Notable Activity_SMS Sent

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.**

ANS

The cutoff value should be minimized to include more number of leads which are predicted yes. So , how much we should minimize ? for that we should consider sensitivity and specificity view, the trade off between accuracy , sensitivity and specificity will give us balanced statistics for the model with increased number of leads predicted. Also, X education while making calls to the potential leads can mention that the company would offer referral bonus if the leads would help in making few more students join.

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

As company want to minimize useless phone calls, false positive rate should be minimized. We need maximum True predictive values i.e. precision. Hence, we should be looking at maximizing the precision and choose the best suited threshold from the precision recall curve. Lead Score should be considered and leads having more than 90 score to be considered for phone calls and very limited set of audience who are in dire need and very much interested in the course should be contacted.