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

   Answer: The top 3 variables are, 1. ‘Lead Origin’, 2. ‘What is your current occupation’ and 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?

   Answer: The top 3 categorical/dummy variables are, ‘Lead Origin_Others’, 2. ‘What is your current occupation_Working Professional’ and 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.

   Answer: As the requirement has shifting towards a more aggressive approach to get the leads with even lesser chance of conversion, we should focus more on making positive prediction of a lead being converted, so that, even if we mark a few non-conversion as conversion, we should not mark a single conversion as non-conversion. To achieve this, the threshold has to modified to as close to zero as possible in such a way that we get a score of zero for **‘False Negative’** in confusion matrix.

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: As the requirement has shifting towards a softer approach to get the leads with only high chance of conversion, we should focus more on making positive prediction of a lead not being converted, so that, we do not miss any leads which has no chance of conversion. To achieve this, the threshold has to modified to as close to one as possible in such a way that we get a score of zero for **‘False Positive’** in confusion matrix.