Lead Scoring Case Study Subjective Questions

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

As per the feature selection step performed by RFE to produce top 15 features and then manually picking out the variables based on p-value and VIF score. The parameters having highest absolute coefficients in the resultant model are:

- Lead Origin
- Last Activity
- 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?**

The top Three categorical/dummy variables which increase the probability of lead conversion in the resultant model:

- Lead Origin_Lead Add Form
- Last Activity_Had a Phone Conversation
- Last Notable Activity_Unreachable

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.**

In the given scenario, since the sales team have interns as well to make the calls, we can imply a strategy it increase the true positives and true negatives in our confusion matrix so that even no leads are missed and sales team can now make use of their full resources to check that.

Hence, this will imply technically that the sensitivity of the model should be high and the cutoff point should be selected as such which maximizes that. From our analysis, the cutoff should be kept below 0.5 to obtain better results.

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.**

Here the given scenario is that the target is already achieved and now since the team needs to focus on new work, they will require to only approach the most important or the true positive leads which will convert, hence in the model we will require to increase the value of TRUE POSITIVES and lower the value for FALSE POSITIVE.

The above scenario can be achieved by considering PRECESION that needs to be increased. Hence, the cutoff needs to be selected which increases that, therefore a cutoff value greater that 0.5 will help here to achieve the results.