1. Which are the top three variables in your model which contribute most towards the probability of a lead getting converted?
   A: The top three variables in our model which contribute most towards the probability of a lead getting converted are Total Time Spent on Website, Lead Origin_Lead Add Form, Last Notable Activity_Modified with absolute z values of 27.296, 18.782 and 18.008 respectively. Their coefficients are 0.9343, 3.4568 and -1.3824.

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?
   A: All the variables in our model are binary variables except Total time spent on website. Apart from the two categorical variables Lead_origin_Lead Add Form and Last Notable activity Modified we specified above, we need to focus on people who have clicked their email links in their last notable activity and who are working professionals to increase the probability of lead 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.
   A: Here the objective is to maximize the true positives as much as we can so that the interns have a complete list of potential leads. This list of a higher potential leads can be obtained by focusing on a model which has a high sensitivity (True positive rate). We achieve this by choosing a threshold for our model that gives us a higher TPR at the expense of other performance metrics.

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.
   A: In this situation, our objective is to minimize the False positive rate, so that we don't mistakenly flag a non-lead as a hot lead. To achieve this, we can build a classification model which focuses on maximizing specificity. This is so because False positive rate = 1- specificity and to minimize FPR we need to maximize Specificity. We will choose a reasonable threshold which maximizes specificity at the expense of other performance metrics.