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

Ans: The overall top three variables are ‘Tags_Closed by Horizzon', 'Tags_Lost to EINS' and 'Tags_Will revert after reading the email'

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: The top 3 categorical/dummy variables are ‘Tags_Closed by Horizzon', 'Tags_Lost to EINS' and 'Tags_Will revert after reading the email'

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: A good strategy is we will decrease the threshold or optimal cutoff, so that we get more positive values thus the sensitivity is increased and specificity is decreased. Example: if threshold is 0.7 then $P(x) < 0.4$, where $y = 0$ and $P(x) > 0.4$, where $y = 1$. Means due to low cutoff, probability of x (student) will be fall more into class 1 (conversion)

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: A strategy will be that we will increase the threshold or ideal cutoff. So, when we increased the threshold, we get more negative values than the positive values thus specificity is increased and sensitivity is decreased. Example: if threshold is 0.7 then $P(x) < 0.7$, where $y = 0$ and $P(x) > 0.7$, where $y = 1$. Means due to high cutoff, probability of x (student) will be fall more into class 0 (non-conversion)