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

Sol: The top Three variables that contribute to result are:

- Total Visits
- Total Time Spent on Website.
- Lead Origin lead add form
- 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?

Sol: The Top 3 Categorical to be focused are:

- Lead Origin lead add form
- Lead Source olark chat
- Lead Source welingak website
- 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.

Sol: Phone calls must be done to people if:

- They are working professional.
- Their last activity is through SMS or through Olark Chat Conversation.
- They coming back to website.
- 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.

Sol: A good strategy will be:

- To focus on narrow set of lead audience (discarding lower conversion probable leads)
- Technically, we can generate this new set of leads by altering (moving up) the value of cut off so as to discard lower conversion rate prabable leads from our Logistic Regression Model
- Doing so, we will be doing minimal effort and still be getting fair conversions.