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

- The following are the top three variable which contributes towards probability of a lead conversion –
    I. Tags_Lost to EINS
    II. Tags_Closed by Horizzon
    III. Lead Quality_Worst

  Reasons – The coefficient for above are (~9.58), (~8.56), and [~ (-)3.94]. These variables have more absolute weightage then rest of the variables

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?

- To increase the probability of lead conversion, we should only consider variables which contributes positively towards the probability of the conversion. The feature which have high positive coefficient are –
    I. Tags_Lost to EINS
    II. Tags_Closed by Horizzon
    III. Tags_Will revert after reading the email

  Reasons – The coefficient for above are (~9.58), (~8.56), and [~ (-)3.94]. This variables have more weightage.

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.

- The sales team can suggest interns to follow up with the customers who are tagged as “Closed by Horizzon”, “Lost to EINS”, and “Will revert after reading the email” for further updates as this category has high probability for conversion
- Target those leads which are sourced from Welingak Website. The interns can also try to gain more leads from Welingak Website.
- The interns can sort the list of the customers whose has sent the SMS. This list will help the sales team to concentrate smaller subset of the dataset.
  Reason –
    - All the variables listed in the above points have positive coefficient, as per the logistic regression model.
    - If properly worked on the variables, the chances of converting the leads will increase.
    - Hence the strategies considered the above variables.

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.

- With the help of the model developed, the sales team can ignore those leads which have low probability of conversion.
- If considering the features of the model, the sales team can ignore leads with following classes of Tags –
  I. Ringing
  II. Switched off
- Can ignore leads with following classes of leads quality –
  I. Not Sure
  II. Worst
- Can ignore leads with following classes of last Notable Activity –
  I. Modified
  II. Olark Chat Conversation.

  Reason –
  - The above variables impact negatively on the probabilities. Hence it is suggested to ignore those customers which have the above attributes.

For reference, following is the tables of parameters produced by the logistic regression model –

Do Not Email	-1.1805
Lead Origin_Lead Add Form	0.908052
Lead Source_Welingak Website	3.21816
Last Activity_SMS Sent	1.927033
Tags_Busy	3.649486
Tags_Closed by Horizzon	8.555901
Tags_Lost to EINS	9.578632
Tags_Ringing	-1.77138
Tags_Will revert after reading the email	3.831727
Tags_switched off	-2.33668
Lead Quality_Not Sure	-3.47923
Lead Quality_Worst	-3.94368
Last Notable Activity_Modified	-1.68208
Last Notable Activity_Olark Chat Conversation	-1.30494