1. Which are the top three variables in your model which contribute most towards the probability of a lead getting converted?
   The top 3 variables are:
   a. **Lead Source:** All lead source variables are showing positive coefficients, this means that this is an important contributor for lead conversion. This shows that lead provided by some sources are much more reliable, and show high conversion.
   b. **Tags:** This variable is showing large positive coefficient and large negative coefficients in the model. This is an important contributor. This shows that lead tagging is also a reliable indicator, and show high conversion.
   c. **What is your current occupation:** This also a good indicator of lead conversion. Working professionals show higher conversions but students and unemployed are also important leads.

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 3 categorical / dummy variables are:
   a. **Lead Source_Welingak Website:** With coefficient 5.0336 this is the top contributor. This means, leads received from Welingak website are hot leads.
   b. **Tags_Ringing:** This is the top negative contributor with a score of -3.9399. This means, those leads which do not pick the call and leave the call ringing are least likely to take the course.
   c. **Tags_Will revert after reading the email:** This is again a positive contributor, with a coefficient of 3.7093. This may mean that a lead interested in reading the email can be a hot lead.

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 company wants to connect with all the potential leads. To do this, they can take the cutoff probability (0.35) or Lead Score (35) as the limit. In this case, they will be able to utilize the maximum number of leads available.
   In terms of variables, based on the outcomes of our LR model, we can ask intern to focus more on the following leads:
   a. whose lead source is Welingak website or a reference.
   b. that are working professionals.
   c. that are tagged ‘Will revert after reading the email’.
   d. whose last activity was ‘SMS sent’.

   Deprioritize the following leads:
   a. that are tagged ‘Ringing’ or ‘Interested in other courses’.
   b. whose lead quality is marked as ‘Worst’.

c. whose Asymmetrique Activity Index is marked as ‘Low’.

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.
   Now, the company wants to connect with the extremely potential leads. To do this, they can take the cutoff probability (0.9) or Lead Score (90) as the limit. In this case, they will be able to minimize the number of false positive leads.

   In terms of variables, we can ask cut down on the number of calls by focusing on only those clients which satisfy all the 7 points (4 positives and 3 negatives) discussed above in the answer to question 3.