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

    1) **Total Time Spent on Website: True –** This has a positive correlation, higher the time spent on the website, higher the chances of a lead generated getting converted into a customer. Thus, pushing more custom offers would easily convert more such leads.
    2) **Lead Source_Others –** This variable also has a positive co-relation, if the lead source is mentioned as others then we need to keep an eye on such leads. They should not be let to go away from the website from shopping. Generating custom messages about relevant products would help in such conversions.
    3) **What is your current occupation_Working Professional –** If the lead is a Working professional, he or she might not be interested in courses of other categories such as students. Thus, preference wise suggestion could improve the conversion rate and increasing the customer base.

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?
    - Lead Source Reference
    - Last Notable Activity SMS Sent
    - Last Notable Activity Email Opened
    - Tags Will revert after reading email

    Be it the top 3 variables or the dummy variables, Lead Source has a higher influence on the lead conversion rate. Thus focus on the lead source with keeping a check on the last activity could improve the leads conversion rate.

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 good approach would be to focus on the maximizing the Recall Factor. Recall is the metric that measures the proportion of actual positives (potential leads that should be converted) that are correctly identified by the model.
    We can also do the following,
    - Lower the decision threshold of the model. This means that the model will classify more customers as potential leads (predicted as 1). While this might reduce precision slightly (more false positives), it will ensure that the model captures most of the true leads.
    - After predicting potential leads, the sales team could manually review the cases where the model is less confident (those closer to the decision threshold). This ensures that no potential lead is left out due to model uncertainty.

- Deploy the interns to reach out to all predicted leads, prioritizing those with the highest predicted probability. This ensures that high-potential leads are contacted first.

By focusing on maximizing recall, X Education can ensure that they reach out to as many potential leads as possible, leveraging their intern resources effectively during this critical hiring period.

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 that the target has been achieved by the company, they should focus on the leads that would convert for sure. Thus, increasing the threshold limit of potential leads would be an advantage. The sales team can make lesser calls with higher conversion rate. Also, focus on the leads with current occupation as working professions will have better chances of conversion, as unemployed or students might not be able to invest a good sum on the course.