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

   The top three variables that contribute most towards the probability of a lead getting converted are:

- Total Time Spent on Website (Coefficient: 3.3576)
- Lead Origin_Lead Add Form (Coefficient: 2.7748)
- What is your current occupation_Working Professional (Coefficient: 1.2066)

   These coefficients indicate that leads who spend more time on the website, come through the Lead Add Form, and are working professionals have a higher probability of getting converted.

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 that should be focused on the most in order to increase the probability of lead conversion are:

- Lead Origin_Lead Add Form (Coefficient: 2.7748)
- What is your current occupation_Working Professional (Coefficient: 1.2066)
- Lead Origin_Landing Page Submission (Coefficient: -0.7772)

   These variables indicate that leads who come through the Lead Add Form and are working professionals have a higher probability of getting converted. On the other hand, leads who come through the Landing Page Submission have a lower probability of getting converted. Therefore, the focus should be on optimizing the Lead Add Form and targeting working professionals to increase the probability of lead conversion. Additionally, efforts can be made to improve the landing page experience to increase the probability of conversion for leads coming through the Landing Page Submission.

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.**

   One strategy that X Education can employ during this phase is to focus on increasing the recall of their model. Recall is the proportion of actual positive cases that are correctly identified by the model as positive. In this case, the positive cases are the potential leads who have been predicted as 1 by the model.

By increasing the recall, X Education can identify a larger proportion of potential leads who are likely to convert, and focus their phone calls and efforts on these individuals. This can lead to a higher conversion rate and more efficient use of resources.

To increase recall, X Education can consider adjusting the model's decision threshold. By lowering the threshold, the model will classify more cases as positive, potentially capturing more of the potential leads.

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.**

   During this time, the company can adjust the threshold for predicting positive leads to be higher, which would result in a lower number of predicted positive leads. This can be done by analysing the trade-off between precision and recall. By setting a higher threshold, the model is less likely to predict a lead as positive unless it is very confident about it, which would result in a higher precision (lower false positive rate) but lower recall (higher false negative rate).

   To implement this strategy, the company can analyse the ROC curve and determine the threshold that provides the best balance between precision and recall, based on their specific goals and priorities. By setting this threshold, the company can ensure that the sales team focuses only on the most promising leads, while minimizing the number of useless phone calls.