<u>Subjective Questions</u>

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

   Response:
   The top three variables in our model contributing most to the probability of a lead conversion can be determined by analyzing the absolute values of the coefficients in the logistic regression model. Higher absolute coefficient values indicate stronger influence on the dependent variable (lead conversion probability).

   Coefficients are:
   1. Tags_Closed by Horizzon
   2. Tags_Lost to EINS
   3. Tags_Will revert after reading the email

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?

   Response:

   To identify top categorical/dummy variables:

   Filter the variables that are categorical (typically encoded as dummy variables).
   Repeat the process of sorting by the absolute coefficient values for these variables.

   Top 3 Categorical Variables:
   1. Tags
   2. Lead Source
   3. Country

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.

   Response:
   During periods with aggressive lead conversion goals:

Lower the classification threshold: By reducing the threshold for classifying a lead as 1 (likely to convert), the model will predict more leads as 1. This increases recall but may reduce precision .Presently we have considered our threshold to be 0.3. This can be moved down to 0.2.

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.

Response:
During times when the company wants to avoid unnecessary phone calls:

Increase the classification threshold: By increasing the threshold, the model will predict fewer leads as 1, focusing only on those with a high probability of conversion. This increases precision but may reduce recall.

Presently we have considered our threshold to be 0.3. This can be moved down to 0.8.