Subjective Questions

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

   The top three variables in the model are:
   - Lead source
   - Occupation
   - Last origin

2. What are the top three categorical/dummy variables in the model which should be focused the most on in order to increase the probability of lead conversion?

   The top three categorical/dummy variables in the model are:
   - Lead sources from "Welingak Websites"
   - Lead Origin from "Lead Add Form"
   - Leads who have occupation as "Working Professionals"

3. X Education has a period of two months every year, that during that time, they hire some interns. The sales team, in particular, has around 10 interns working there. So during this phase, they wish to make the lead conversion rate increase aggressively. Therefore, they want almost all the potential leads (i.e., the customers who have been predicted as 1 by the model) to be converted. Hence, they can make phone calls to as many leads as possible. Suggest a good strategy that they should employ at this phase.

   Sensitivity with respect to the model can be defined as the ratio of total number of actual conversions correctly predicted to the total of number of actual conversions. For a particular model, as Sensitivity increase, Specificity decreases and vice versa. Different values of the sensitivity and specificity can be achieved for the same model by changing the Conversion Probability cutoff threshold value.

   When the probability thresholds are very low, the sensitivity is very high, and specificity is very low. High sensitivity implies that the model will correctly identify almost all leads who are likely to convert. It will do that by over estimating the Conversion likelihood, i.e., it will misclassify some non-conversion cases as conversions.

   Since X Education has more employees for 2 months, they wish to make the lead conversion rate increase significantly, we can choose a lower threshold value for Conversion Probability. This will ensure the Sensitivity rating to be high which in turn will make sure almost leads who are likely to convert are identified correctly and the agents can make phone calls to as many people as possible.

   The following leads must be contacted if:

- They have occupation as working professionals
- They visit and spend considerable amount of time on the website
- Lead source is either from Welingak website or References or Olark Chat conversation
- Lead origin is "Add Form"

4. Similarly, at times, the company reaches its target for a quarter before the deadline. During that time, the company wants the sales team to focus on some new work as well. So during that time, the company avoids making 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 that stage.

To minimize the rate of useless phone calls, the company can:
- Choose a higher threshold value for Conversion Probability. This ensures that all the leads contacted have a higher conversion probability.
- Call or send automated emails and SMS to target leads with high conversion probability (Hot leads).