Assignment Subjective Questions:

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

   Ans: The top three variables in our model which contribute most towards the probability of a lead getting converted are:

   a. Total Time Spent
   b. Lead_Origin_Lead Add Form
   c. LeadSource_Welingak Website

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?**

   Ans: 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 are:

   a. Lead_Origin_Lead Add Form
   b. LeadSource_Welingak Website
   c. Current_Occupation_Working Professional

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.**

   Ans: Sensitivity with respect to our model can be defined as "the ratio of total number of actual Conversions correctly predicted to the total no of actual Conversions".
   Similarly, **Specificity** can be defined as "the ratio of total no of actual non-Conversions correctly predicted to the total number of actual non-Conversions".

➢ For a particular model, as one increases, the other 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.

For our model, the below graph shows how the Sensitivity and Specificity rating changes with change in the threshold value:
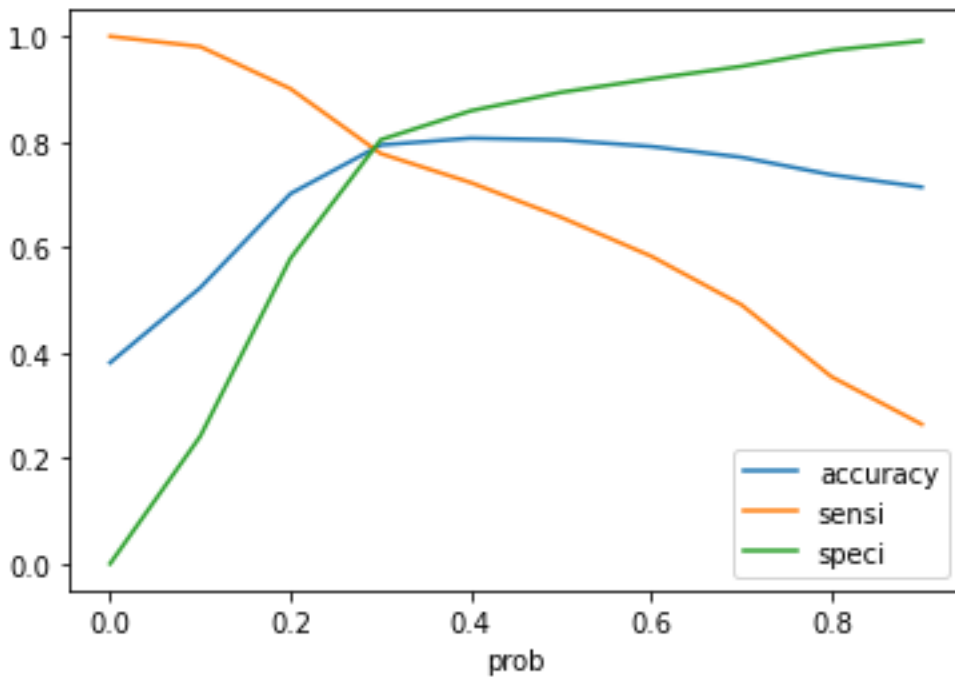
➢ When the probability thresholds are very low, the sensitivity is very high and specificity is very low. Similarly, for larger probability thresholds, the sensitivity values are very low but the specificity values are very high.

➢ High sensitivity implies that our 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.

➢ Now, since X Education has more man-power for these 2 months and they wish to make the lead conversion more aggressive by wanting

almost all of the potential leads, **we can choose a lower threshold value for Conversion Probability.**

➢ This will ensure the Sensitivity rating is very high which in turn will make sure almost all leads who are likely to Convert are identified correctly and the agents can make phone calls to as much of such people as possible.

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.**

Ans: Similar to the above logic and context from the previous question, High Specificity implies that our model will correctly identify almost all leads who are not likely to Convert. It will do that at the cost of losing out some low Conversion rate risky leads to the competition, i.e., it will misclassify some Conversion cases as non-Conversions.

Therefore, since X Education has already reached its target for a quarter and doesn't want to make phone calls unless it's extremely necessary, i.e., they want to minimize the rate of useless phone calls, we can choose a **higher threshold value for Conversion Probability.**

This will ensure the Specificity rating is very high, which in turn will make sure almost all leads who are on the brink of the probability of getting Converted or not are not selected. As a result, the agents won't have to make unnecessary phone calls and can focus on some new work.