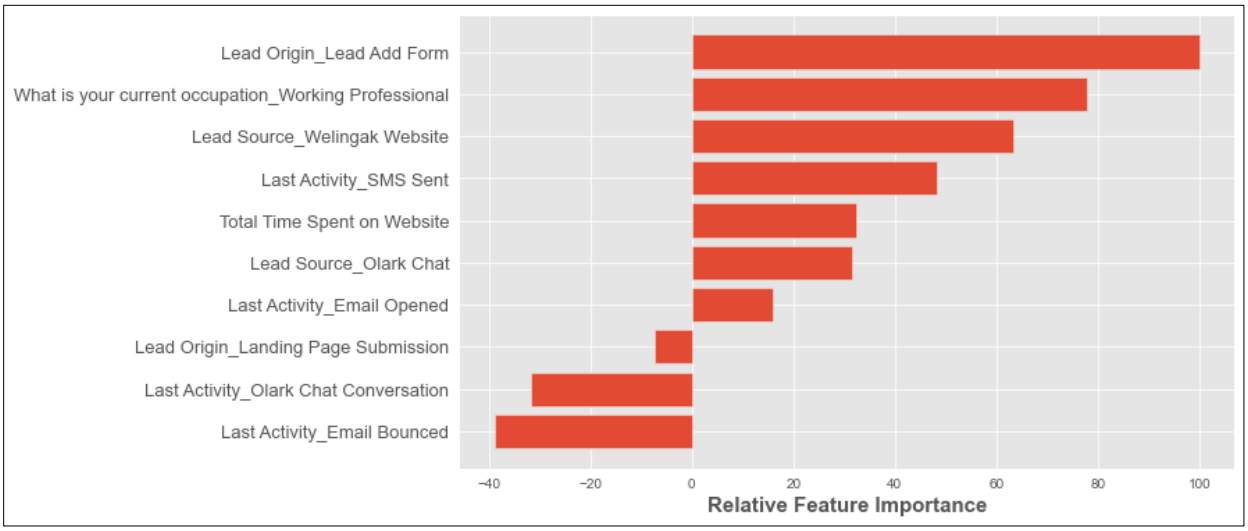
Subjective Questions

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

A. Based our model, 7 of the 10 predictors belong to the below three variables:

- Lead Origin: Lead add Form & Landing Page Submission
- Lead Source: Welingak Website & Olark Chat.
- Last Activity: SMS Sent, Email Opened, Email Bounced



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?**

A. Based on the coefficient values, the following are the top three categorical/dummy variables that should be focused on the most in order to increase the probability of lead conversion:

- Lead Origin_Lead Add From
- What is your current occupation_Working Professional
- Lead Source_Welingak Website

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**

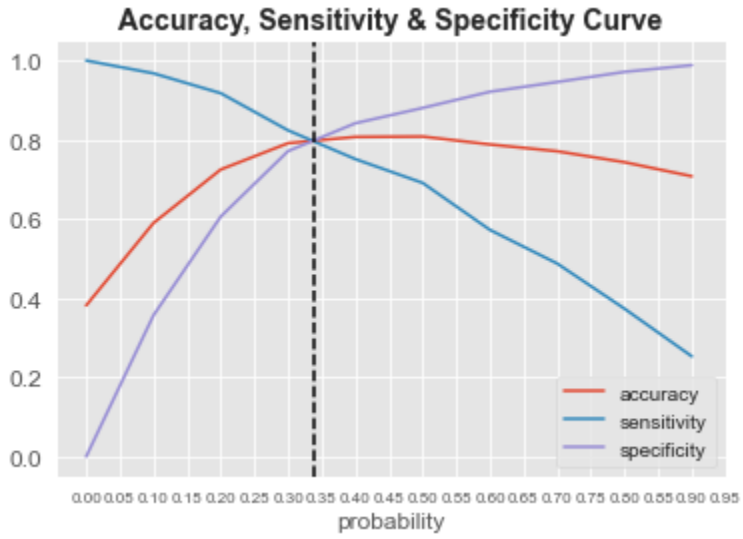
of such people as possible. Suggest a good strategy they should employ at this stage.

A. Here, the concept of sensitivity is at play.

- In the given situation, we will need a tweak the model to increase its sensitivity because high sensitivity will mean that our model will correctly predict almost all leads who are likely to convert.
- To achieve high sensitivity, we need to choose a low probability threshold value.

Below is an explanation:

- $\boldsymbol{Sensitivity} = \boldsymbol{True\ Positives}/ (\boldsymbol{True\ Positives} + \boldsymbol{False\ Negatives})$
- With respect to our model, sensitivity can be defined as **"of all the leads that converted, how many were correctly predicted as converted by our model**.
- Different values of sensitivity can be achieved for the model by changing the probability cutoff threshold for lead conversion.
- For our model, the below graph shows changes in Sensitivity, Specificity and Accuracy with change in the probability threshold:



- As we can see, sensitivity decreases with every increase in the cut-off threshold.
- However, we also need to be cognizant that sensitivity and precision are inversely correlated hence when sensitivity is increased the precision would decline and vice versa.
- In our case it would lead to our model misclassifying some of the non-converted leads as converted.
- But as the company has extra man-power for two months and wants to make the lead conversion more aggressive by making phone calls to as much potential leads as possible, it is a good strategy to go for high sensitivity.

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.**

   - Here, the concept of specificity is at play.
   - In the given situation, will need a high specificity because high specificity will mean that our model will correctly predict almost all leads who are not likely to convert. To achieve high specificity, we need to choose a high threshold value.

   <u>Below is an explanation:</u>

   - $\boldsymbol{Specificity} = \boldsymbol{True\ Negatives}/ (\boldsymbol{True\ Negatives} + \boldsymbol{False\ Positives})$
   - With respect to our model, specificity can be defined as "**of all the people who did not convert, how many did our model correctly predict".**
   - As seen in the graph above, we can see that the specificity increases as the threshold increases.
   - However, increasing the specificity may lead to misclassifying some of the converted leads as non-converted.
   - As the company has already reached its target for a quarter and doesn't want to make phone calls unless it's extremely necessary, it is a good strategy to go for high specificity.
   - It will ensure that the phone calls are only made to customers who have a very high probability of conversion.