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

A1) We have constructed a graph illustrating the relative significance of different variables based on the coefficient values derived from our model.

As per the graph, following are the top 3 variables which contribute the most towards the probability of a lead getting converted-

- **Lead Origin_Lead Add Form**
- **Last Activity_Had a Phone Conversation**
- **What is your current occupation_Working Professional**
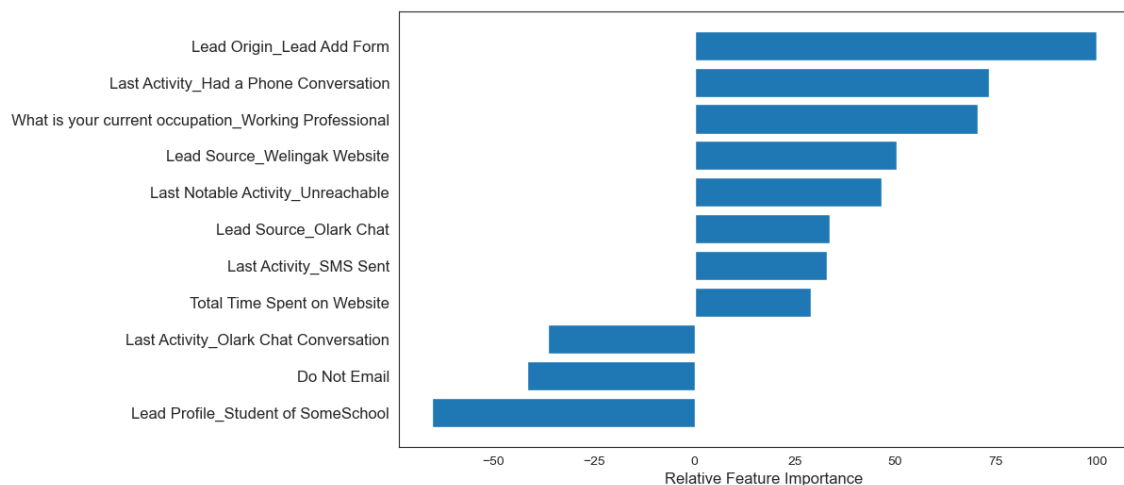
2. What are the top 3 categorical/dummy variables in the model which should be focused the most on in order to increase the probability of lead conversion?

A2) According to the graph above, the top 3 variables are categorical/dummy variables. Hence, the answer remains the same.

As per the graph, following 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 -

- **Lead Origin_Lead Add Form**
- **Last Activity_Had a Phone Conversation**
- **What is your 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.

A3) In our model, the graph below depicts variations in Sensitivity, Specificity, and Accuracy in response to alterations in the threshold.
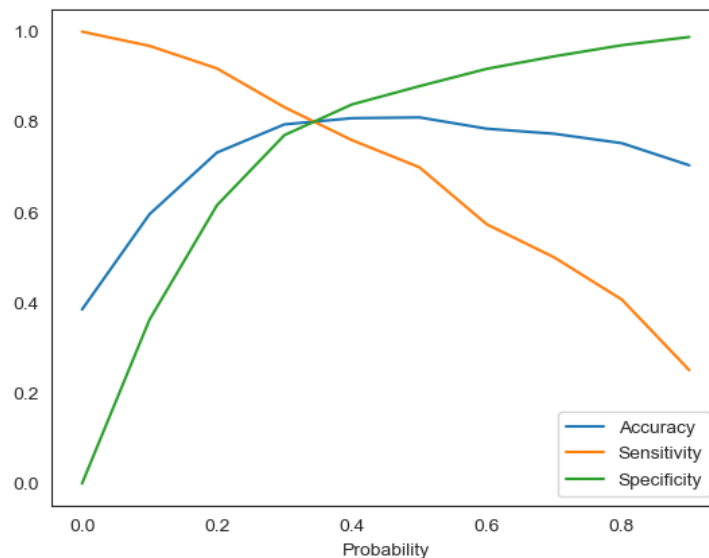
Concerning our model, sensitivity is characterized as the proportion of accurately predicted actual conversions to the total number of actual conversions.

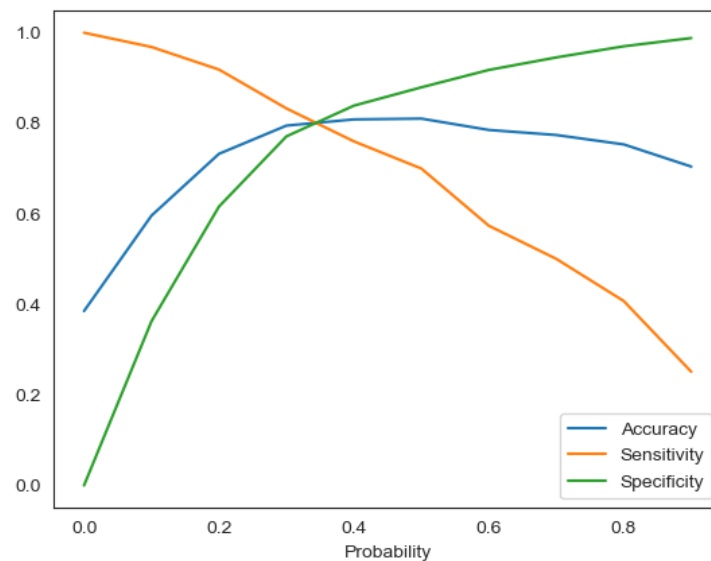$$Sensitivity = True\ Positives\ /\ (True\ Positives + False\ Negatives)$$

As observed, sensitivity diminishes with an increase in the threshold. In this scenario, a high sensitivity is crucial, as it ensures that the model accurately identifies nearly all leads likely to convert.

Given that the company has surplus manpower for the next two months and aims to intensify lead conversion efforts by reaching out to as many potential leads as feasible, opting for high sensitivity is a strategic choice. To attain high sensitivity, **selecting a low threshold value is essential**.

Hence, the company may contact all the leads which have a conversion probability (value = 1) **under a cut off 0.3 to 0.4 for better results.**

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.

A4) In our model, the graph below depicts variations in Sensitivity, Specificity, and Accuracy in response to alterations in the threshold.

In the context of our model, specificity is defined as the count of accurately predicted non-conversions out of the total number of actual non-conversions.

$$Specificity = True\ Negatives\ /\ (True\ Negatives + False\ Positives)$$

From the above graph, it is evident that **specificity rises with an increase in the threshold**. In this particular scenario, a high specificity is desirable, as it indicates that the model will accurately predict almost all leads that are not likely to convert.

To attain high specificity, **opting for a high threshold value** is necessary.

Hence, the company may contact all the leads which have a conversion probability (value = 1) **above a cut off 0.7 or 0.8 for better results**.