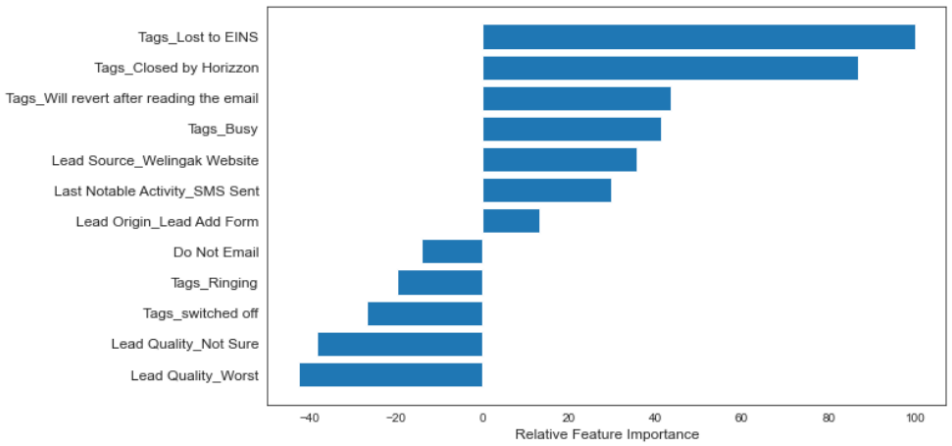
1. Which are the top three variables in your model which contribute most towards the probability of a lead getting converted?
   Ans –



   The three variables that have the greatest effect on the probability of a lead converting, in decreasing order, are:
   - **I. Tags_Lost to EINS**
   - **II. Tags_Closed by Horizzon**
   - **III. Tags_Will revert after reading the email**

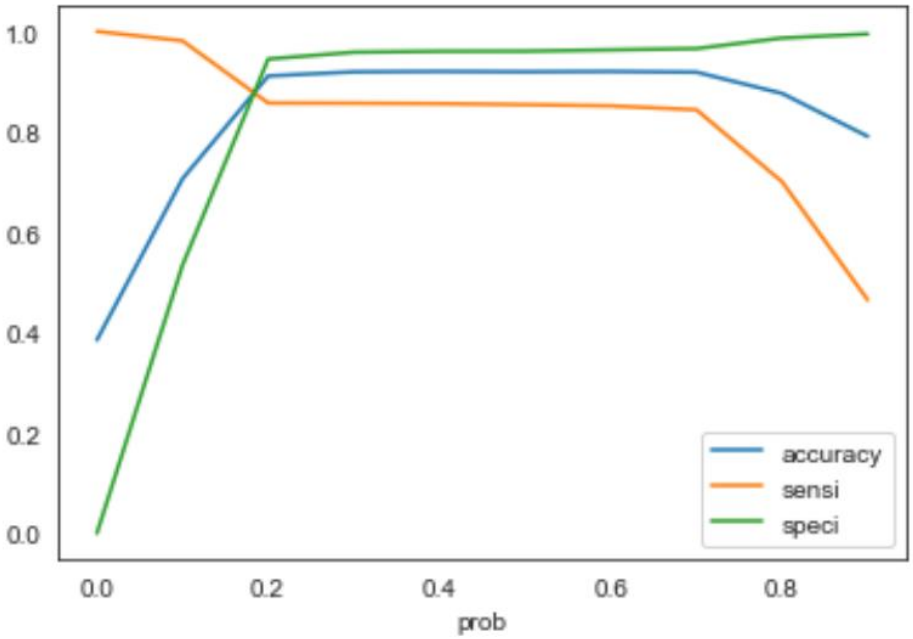
   These variables are derived from the categorical variable 'Tags' and have a positive impact on the likelihood of a lead conversion. These results suggest that the company should prioritize leads with these three tags in order to improve their chances of conversion

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 three categorical/dummy variables that the company should prioritize in order to increase the probability of lead conversion are:
   - **I. Tags_Lost to EINS**
   - **II. Tags_Closed by Horizzon**
   - **III. Tags_Will revert after reading the email**

   These variables are the most important ones to focus on because they have the greatest impact on the probability of lead conversion, according to the model

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 – In this context, sensitivity refers to the number of actual conversions that are accurately predicted by the model, out of the total number of actual conversions.
   $\boldsymbol{Specificity} = \boldsymbol{True\ Negatives}/ (\boldsymbol{True\ Negatives} + \boldsymbol{False\ Positives})$

Sensitivity can be changed by altering the cutoff threshold for the probability of lead conversion.



The graph shows that as the threshold increases, sensitivity decreases. In this case, it is important to have a high sensitivity because it means that the model will correctly predict a high number of leads who are likely to convert.

However, this may also result in the model overestimating and misclassifying some non-conversions as conversions. Since the company has additional resources for two months and wants to aggressively pursue lead conversion by contacting as many potential leads as possible, it is a good strategy to aim for a high sensitivity. To achieve this, a low threshold value should be chosen.

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 - To answer this question, it is important to understand the concept of specificity.
$\boldsymbol{Specificity} = \boldsymbol{True\ Negatives} / (\boldsymbol{True\ Negatives} + \boldsymbol{False\ Positives})$
Specificity is defined as the number of actual non-conversions that are accurately predicted by the model, out of the total number of actual non-conversions.

According to the graph, specificity increases as the threshold increases. In this case, it is important to have a high specificity because it means that the model will correctly predict a high number of leads who are not likely to convert.

However, this may also result in the model misclassifying some conversions as non-conversions. Since the company has already reached its target for the quarter and does

not want to make unnecessary phone calls, it is a good strategy to aim for a high specificity. This will ensure that phone calls are only made to customers who have a very high probability of conversion. To achieve this, a high threshold value should be chosen.