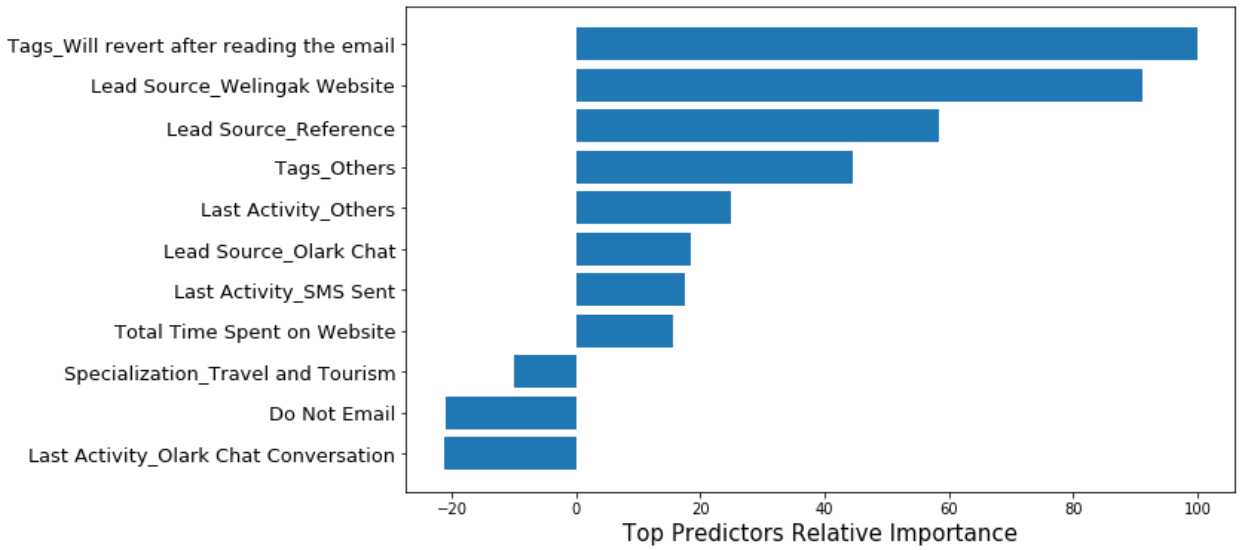
Lead Scoring Case Study Subjective Question

Q1. Which are the top three variables in your model which contribute most towards the probability of a lead getting converted?

A: - The predictors of the final model are shown below based on their importance in converting lead calculated as per the relative coefficient values



Top 3 variables that contribute most towards the probability of a lead getting converted are:

1) Tags_Will revert after reading the email
2) Lead Source_Welingak Website
3) Lead Source_Reference

Q2. 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?

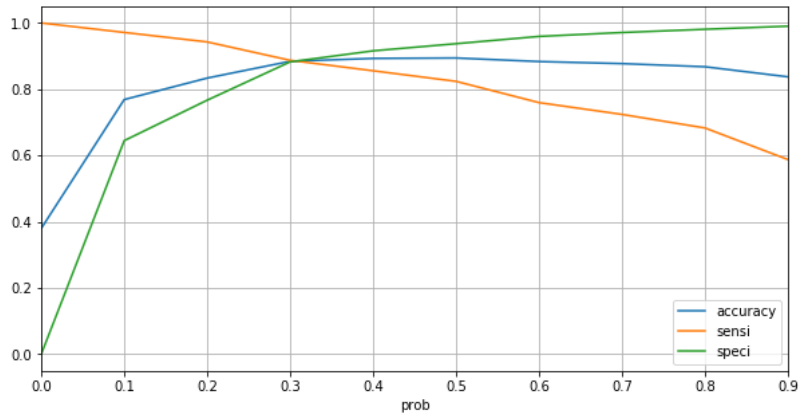
A: - As per the diagram shown above, the top 3 categorical/dummy variables to be focused the most on in order to increase the probability of lead conversion are:

1) Tags_Will revert after reading the email
2) Lead Source_Welingak Website
3) Lead Source_Reference

Q3. 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: - Sensitivity is the probability of positive test of actual Conversions. Similarly, Specificity probability of negative test of actual non-Conversions.

For a model, as one increases, the other decreases and vice versa. Different values of the sensitivity and specificity can be determined by changing the Conversion Probability cutoff threshold value. The curve obtained for our model is shown below: -



Now, for more aggressive lead conversion X-Education needs: -

1. **1st strategy** - To choose a lower threshold value of conversion probability for a high sensitivity, to ensure all positive leads are identified correctly.
2. **The 2nd best strategy** - To hold a tight grip on potential lead by focusing on below variables and consider lead score >=40: -
    1) Tags_Will revert after reading the email
    2) Lead Source_Welingak Website
    3) Lead Source_Reference
    4) Tags_Others
    5) Last Activity_Others

Q4. 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.

A: - Following the similar logic from the previous question, High Specificity implies that model will correctly identify almost all leads who are not likely to Convert. It will do that at the cost of losing or misclassifying 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,

1. **1st Strategy** – To choose a higher threshold value for Conversion Probability
2. **The 2nd best strategy** - To focus only on below variables and consider lead score >=95: -
    1) Tags_Will revert after reading the email
    2) Lead Source_Welingak Website