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

Answer:

```
----------------------Feature Importance--------------------
const                                              -1.248649
Do Not Email                                       -1.180501
Lead Origin_Lead Add Form                           0.908052
Lead Source_Welingak Website                        3.218160
Last Activity_SMS Sent                              1.927033
Tags_Busy                                           3.649486
Tags_Closed by Horizzon                             8.555901
Tags_Lost to EINS                                   9.578632
Tags_Ringing                                       -1.771378
Tags_Will revert after reading the email            3.831727
Tags_switched off                                  -2.336683
Lead Quality_Not Sure                              -3.479228
Lead Quality_Worst                                 -3.943680
Last Notable Activity_Modified                     -1.682075
Last Notable Activity_Olark Chat Conversation      -1.304940
dtype: float64
```

As per the above diagram, the top 3 variables that contribute most towards the probability of a lead getting converted are:

- Tags_Lost to EINS
- Tags_Closed by Horizzon
- Tags_Will revert after reading the email

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?**

Answer:

As per the above diagram, the top 3 categorical/dummy variables that contribute the most towards the probability of a lead getting converted are also :

- Tags_Lost to EINS
- Tags_Closed by Horizzon
- Tags_Will revert after reading the email

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.**

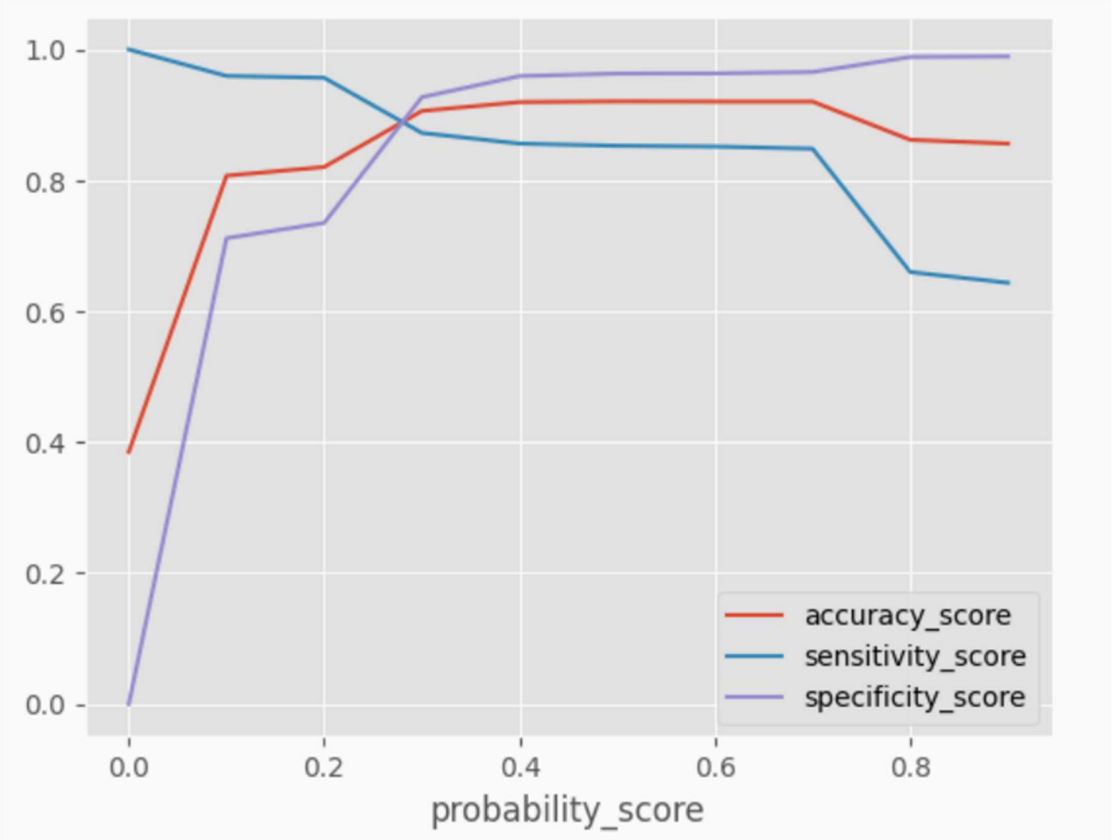
Answer:

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 means that our model accurately identifies almost all convertible leads. It does this by overestimating the likelihood of conversion, i.e. by misclassifying some cases where there is no conversion as conversion.

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 provides a very high sensitivity rating, allowing you to correctly identify almost any potential customer that can convert, and allows agents to call as many of them 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.**

Answer:

Following the similar 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 does this at the cost of giving competitors some risky leads with low conversion rates. That is, it incorrectly classifies some conversions 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. If they want to minimize unwanted phone calls, they can choose a higher threshold 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.