Lead Case Study Questions

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

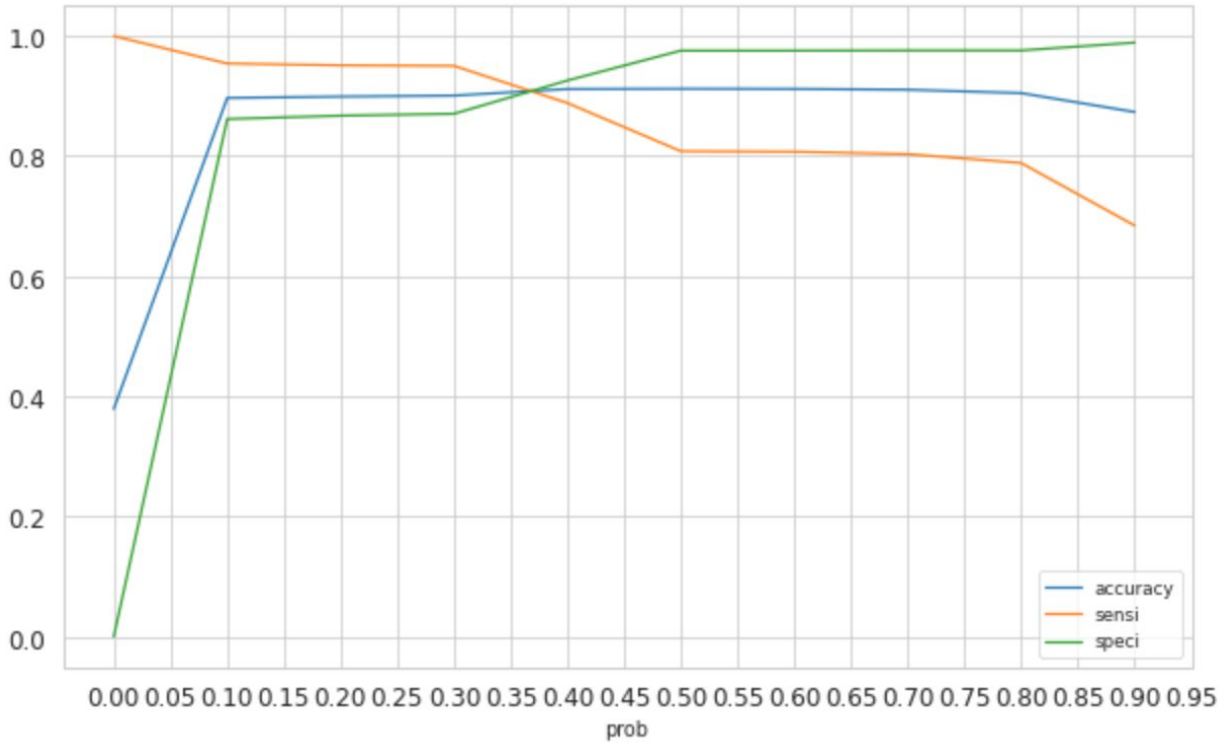
```
Last Activity_sms sent                                37.94
Tags_already a student                               -67.81
Tags_closed by horizzon                               97.45
Tags_interested  in full time mba                    -59.84
Tags_interested in other courses                     -62.38
Tags_invalid number                                  -82.04
Tags_lost to eins                                    100.00
Tags_not doing further education                     -75.74
Tags_opp hangup                                      -52.18
Tags_ringing                                         -83.12
Tags_switched off                                    -96.83
Tags_will revert after reading the email              67.89
Asymmetrique Activity Index_03.low                   -50.50
Lead Quality_worst                                   -43.31
What is your current occupation_unemployed            42.17
What is your current occupation_working professional  45.00
```

   As per the above data, we can see that the top 3 features contributing the most towards the probability of a lead being converted are -
   a. *Tags_lost to eins*
   b. *Tags_closed by horizon*
   c. *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?**

   As in the previous question, we can see the same features that lead to increase in the probability of lead conversion as per the coefficients -
   a. *Tags_lost to eins*
   b. *Tags_closed by horizon*
   c. *Tags_will revert after reading the email*

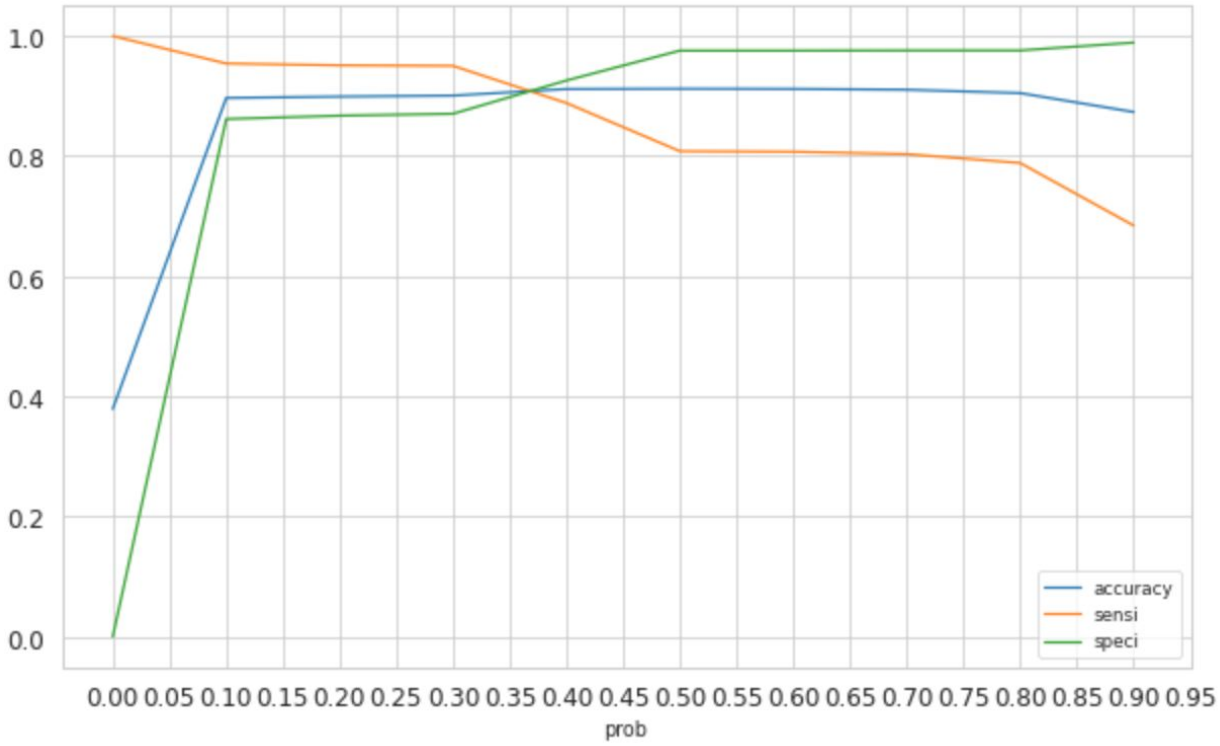
3. **X Education has a period of 2 months every year during which they hire some interns. The sales team, in particular, has around 10 interns allotted to them. So during this phase, they wish to make the lead conversion more aggressive. So they want almost all of the potential leads (i.e. the customers 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.**



Now as per the condition presented in the question, since we are ready to apply a good workforce in the next 2 months, we can target variables that make a high impact on the conversion rate. Also we can reduce the threshold value to know what is the target we can achieve. 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 implies that our model will correctly identify almost all leads who are likely to Convert. It also means it will do over estimating of identifying non-conversion leads as conversion leads. Now this is possible as there is more man-power with the team for the 2 months and they wish to make lead conversion more successful, so they can invest in these factors.
Sensitivity in our case will be the ratio of the number of conversions predicted to the number of conversions we actually had. Now as we lower the threshold, the sensitivity is higher, which means the estimated conversion count is higher , which is an ideal case. Since the sensitivity is high , this will ensure more conversion of leads.

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.**



Now as in the previous problem, we took sensitivity into account as we wanted more leads conversion, but the scenario is different. We have met the targets already and thus need to reduce the expenses on unnecessary phone calls. Now we can take into account the specificity value here.
High specificity means we will correctly identify most of the leads who are not going to convert, but this will come at a cost of losing out some leads who have low probability of getting converted. So this can be identified if we increase the threshold value. It means if we increase the threshold we receive high specificity value , that is we know which all leads will not get converted and don't invest the money there in the unnecessary phone calls and can do something new to achieve the targets in future.