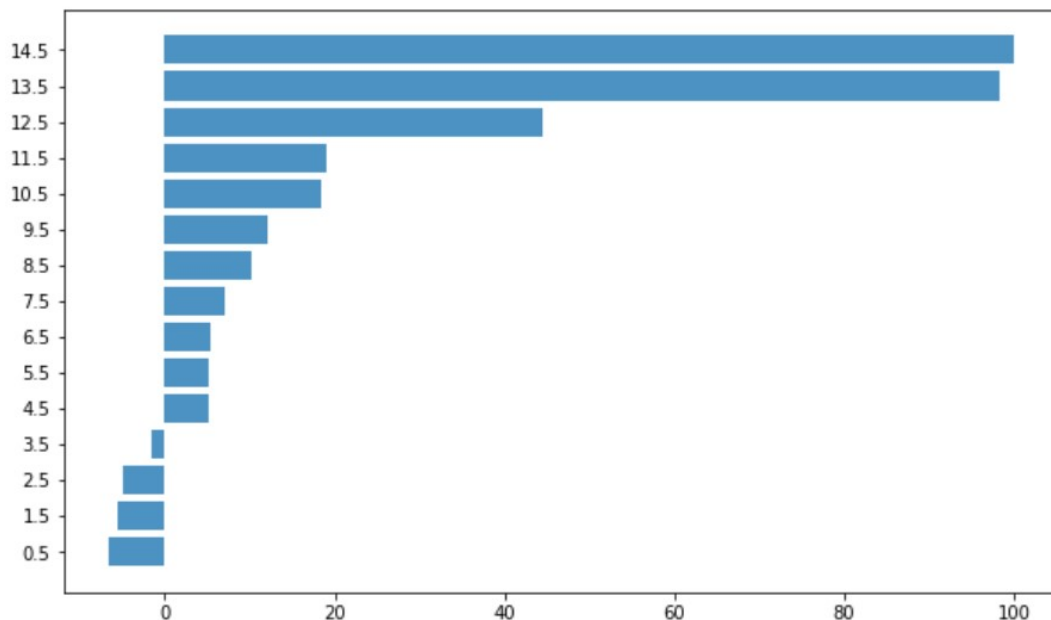


1. Which are the top three variables in your model that contribute most towards the probability of a lead getting converted?

Ans-The features used to build the model have been illustrated based on their coefficient values.



```
sorted_idx = np.argsort(feature_importance,kind='quicksort',order='list of str')
sorted_idx
```

```
: TotalVisits 5
Total Time Spent on Website 10
Lead Origin_Lead Add Form 9
Lead Source_Olark Chat 14
Lead Source_Welingak Website 6
Do Not Email_Yes 7
Last Activity_Had a Phone Conversation 11
Last Activity_SMS Sent 3
What is your current occupation_Housewife 4
What is your current occupation_Student 13
What is your current occupation_Unemployed 2
What is your current occupation_Working Professional 1
Last Notable Activity_Had a Phone Conversation 0
Last Notable Activity_Unreachable 8
Specialization_Select 12
dtype: int64
```

From the above plot it is seen that following 3 variables has the highest effect on lead conversion-

1. Last Notable Activity_Had a Phone Conversation
2. What is your current occupation_Housewife
3. TotalVisits

2. What are the top 3 categorical/dummy variables in the model which get maximum focus in order to increase the probability of lead conversion?

Ans- From the above plot it is seen that following 3 dummy variables has the highest effect on lead conversion-

1. Last Notable Activity_Had a Phone Conversation
2. What is your current occupation_Housewife
3. TotalVisits

3. X Education has a period of 2 months every year during which they hire few 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-Specificity is ratio of total no of actual non-Conversions correctly projected to the total number of definite non-Conversions.

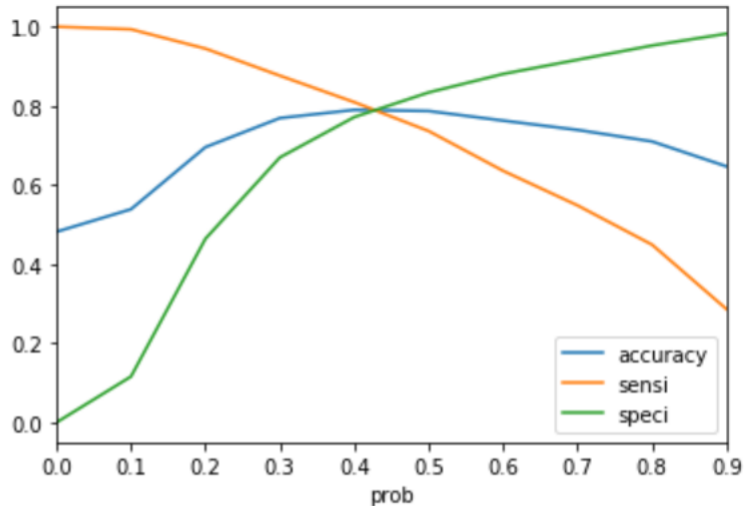
In a specific model, it is generally inversely proportional that is, as one increases, the other decreases and vice versa.

Different values of the sensitivity and specificity can be calculated by changing the cutoff threshold value.

The below graph shows how the sensitivity and specificity rating changes with change in the threshold value:

▶ ##plotting

```
prob_cutoff_df.plot.line(x='prob', y=['accuracy', 'sensi', 'speci'])  
plt.show()
```



Here when probability thresholds are very low, sensitivity value is very high and specificity is very low. Likewise, for higher probability thresholds, sensitivity values are very low but the specificity values are very high. High sensitivity indicates that the model will accurately identify nearly all leads that are likely to convert.

Here, as X Education has more man-power for 2 given months and they need to make the lead conversion more aggressive by needing nearly all of the potential leads, which can be chosen by a lower threshold value for conversion probability. This will guarantee sensitivity rating that is very high which consequently will identify almost all leads who are expected to convert will be identified correctly, so that the agents make phone calls to as much potential customers 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.

Ans- Here high specificity implies that the model will correctly be identified nearly all leads that are not likely to convert. It will do that at the cost of missing some low conversion rate risky leads to the competitors. Hence, as X Education has already reached its target for this quarter and want to make only useful calls i.e., they want to minimize the rate of useless phone calls, we can choose a higher threshold value for conversion probability. This will ensure specificity rating is very high, which in turn will make sure almost all leads getting converted or not are not selected. Thus, the agents won't have to make any needless phone calls and can concentrate on some new work.