

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

Ans:

Top 3 features in final model which contribute most towards the probability of deal getting converted are:

```
What is your current occupation
```

```
Last Notable Activity
```

```
Last Activity
```

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:

Top 3 categorical/dummy variables in final model which contribute most towards the probability of deal getting converted are:

```
'Lead Source_Welingak Website',
```

```
'Last Notable Activity_Had a Phone Conversation',
```

```
'Lead Origin_Lead Add Form'
```

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:

My strategy In this case the Recall should be high and precision can be compromised. And using the graph created between probability, accuracy, recall, precision we can choose cut-off point of 0.2 to high Recall and considerable precision.

Model performance at this cut-off:

```
test Scores
```

```
Accuracy: 0.7493680028891296
```

```
Recall: 0.9204119850187266
```

```
Precision: 0.617462311557789
```

```
F1-Score: 0.7390977443609023
```

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:

My strategy In this case the precision should be high and Recall can be compromised. And using the graph created between probability, accuracy, recall, precision we can choose cut-off point of 0.5 to high Recall and considerable precision

Model performance at this cut-off:

test Scores

Accuracy: 0.8057060310581438
Recall: 0.6947565543071161
Precision: 0.7777777777777778
F1-Score: 0.7339268051434225