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

   Answer:

   The top three variables in our model which contributes more towards the probability are,

   1. Lead Source_Welingak Website
   2. Last Activity_Had a Phone Conversation
   3. Lead Origin_Lead Add Form

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:

   Top 3 categorical/dummy variables in the model which should be focused the most on in order to increase the probability are,

   1. Welingak Website (From Lead Source)
   2. Had a Phone Conversation (From Last Activity)
   3. Lead Add Form (From Lead Origin)

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:

```
In [94]: y_train_pred_final['Lead_Score'] = y_train_pred_final.Converted_prob.map( lambda x: round(x*100))

         y_train_pred_final.head()
```

Out[94]:

| | Converted | Converted_prob | Prospect ID | predicted | 0.0 | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 | final_predicted | Lead_Score |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **0** | 0 | 0.302757 | 3009 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 30 |
| **1** | 0 | 0.012231 | 1012 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| **2** | 0 | 0.241549 | 9226 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 24 |
| **3** | 1 | 0.688675 | 4750 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 69 |
| **4** | 1 | 0.302757 | 7987 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 30 |

This is the final prediction output. Which is based on a optimum cut-off value of 0.35. So in order to increase the sales, company should focus on contacting leads with a conversion probability of 1 under a cutoff of 0.35

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:

In order to minimize the rate of useless phone calls, the company may contact all the leads which have a conversion probability (value = 1) under columns 0.5,0.6,0.7,0.8,0.9. Because in that columns, there are leads that actually converted. So there may be a chance to miss out these leads.

| | Converted | Converted_prob | Prospect ID | predicted | 0.0 | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 | final_predicted | Lead_Score |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **0** | 0 | 0.302757 | 3009 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 30 |
| **1** | 0 | 0.012231 | 1012 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| **2** | 0 | 0.241549 | 9226 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 24 |
| **3** | 1 | 0.688675 | 4750 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 69 |
| **4** | 1 | 0.302757 | 7987 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 30 |
| **5** | 1 | 0.688675 | 1281 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 69 |
| **6** | 0 | 0.302757 | 2880 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 30 |
| **7** | 1 | 0.896792 | 4971 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 90 |
| **8** | 1 | 0.302757 | 7536 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 30 |
| **9** | 0 | 0.302757 | 1248 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 30 |
| **10** | 0 | 0.302757 | 1429 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 30 |
| **11** | 0 | 0.074997 | 2178 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 7 |
| **12** | 0 | 0.302757 | 8554 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 30 |
| **13** | 1 | 0.618676 | 5044 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 62 |
| **14** | 1 | 0.302757 | 3475 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 30 |
| **15** | 1 | 0.688675 | 7424 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 69 |
| **16** | 0 | 0.302757 | 421 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 30 |
| **17** | 0 | 0.241549 | 3591 | 0 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 24 |
| **18** | 0 | 0.062193 | 6247 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 6 |
| **19** | 0 | 0.618676 | 7843 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 62 |