SUBJECTIVE QUESTION ANSWERS

Q1) which are the top three variables in your model, which contribute most towards the probability of a lead being converted?

Based on the coefficient values from below screeshot, the following are the top three variables that contribute most towards the probability of a lead getting converted :

a. Lead Source_Welingak Website
b. Last Notable Activity_Unreachable
c. What is your current occupation_Working Professional

	coef
const	0.4242
Total Time Spent on Website	1.0974
Lead Source_Direct Traffic	-1.8221
Lead Source_Google	-1.3003
Lead Source_Organic Search	-1.5290
Lead Source_Reference	2.2568
Lead Source_Referral Sites	-1.6816
Lead Source_Welingak Website	3.9887
Do Not Email_Yes	-1.4230
Last Activity_Had a Phone Conversation	1.8863
Last Activity_Olark Chat Conversation	-1.4742
Last Activity_SMS Sent	1.1943
What is your current occupation_Working Professional	2.3699
Last Notable Activity_Unreachable	2.5080

Q2) 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?

Again, based on the coefficient values from the screen shot in the question above, the following are the top three categorical/dummy variables that should be focused the most in order to increase the probability of lead conversion :

a. Lead Source_Welingak Website

b. Last Notable Activity_Unreachable

c. What is your current occupation_Working Professional

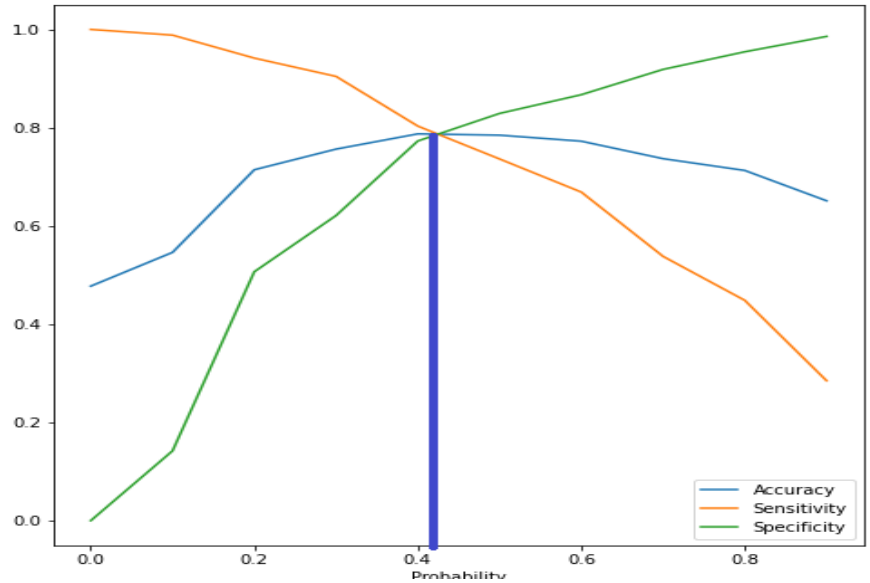
Q3) X Education has a period of 2 months every year during which they hire some interns. The sales team, in particular, has around 10 interns allotted to them. Therefore, during this phase, they wish to make the lead conversion more aggressive. Therefore, they want almost all of the potential leads (i.e. 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.

In the below image, the final prediction is calculated based on optimal cut off value of 0.42.

In order to make the sales aggressive, the company may contact all the leads which have a conversion probability (value = 1) under a cut off 0.4(yellow marks).

```
y_train_pred_final
```

	Converted	Leadscore_Prob	CustID	predicted	leadscore_score	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	final_predicted
0	0	0.356556	1871	0	35.655553	1	1	1	1	0	0	0	0	0	0	1
1	1	0.518751	6885	1	51.875063	1	1	1	1	1	1	0	0	0	0	1
2	0	0.107384	3427	0	10.738368	1	1	0	0	0	0	0	0	0	0	0
3	0	0.882647	1525	1	88.264676	1	1	1	1	1	1	1	1	1	0	1
4	1	0.945666	1475	1	94.566566	1	1	1	1	1	1	1	1	1	1	1
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
4157	1	0.909371	3015	1	90.937100	1	1	1	1	1	1	1	1	1	1	1
4158	0	0.665788	615	1	66.578767	1	1	1	1	1	1	1	0	0	0	1
4159	1	0.410769	110	0	41.076924	1	1	1	1	1	0	0	0	0	0	1
4160	0	0.699573	6308	1	69.957267	1	1	1	1	1	1	1	0	0	0	1
4161	0	0.881768	8788	1	88.176847	1	1	1	1	1	1	1	1	1	0	1



Q4) 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?

In order to minimize the rate of useless phone calls, the company may contact all the leads which have a conversion probability (value = 1 highlighted in yellow colour) under column 0.7. However, the flipside here would be that, we may miss those leads that are actually converted but then the model wrongly predicted them as not converted. (Red highlights in the image below). This should not be a major cause for concern as the target has already be achieved.

```
y_train_pred_final
```

	Converted	Leadscore_Prob	CustID	predicted	leadscore_score	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	final_predicted
0	0	0.356556	1871	0	35.655553	1	1	1	1	0	0	0	0	0	0	1
1	1	0.518751	6885	1	51.875063	1	1	1	1	1	1	0	0	0	0	1
2	0	0.107384	3427	0	10.738368	1	1	0	0	0	0	0	0	0	0	0
3	0	0.882647	1525	1	88.264676	1	1	1	1	1	1	1	1	1	0	1
4	1	0.945666	1475	1	94.566566	1	1	1	1	1	1	1	1	1	1	1
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
4157	1	0.909371	3015	1	90.937100	1	1	1	1	1	1	1	1	1	1	1
4158	0	0.665788	615	1	66.578767	1	1	1	1	1	1	1	0	0	0	1
4159	1	0.410769	110	0	41.076924	1	1	1	1	1	0	0	0	0	0	1
4160	0	0.699573	6308	1	69.957267	1	1	1	1	1	1	1	0	0	0	1
4161	0	0.881768	8788	1	88.176847	1	1	1	1	1	1	1	1	1	0	1