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

   Tags_Lost to EINS', 'Tags_Closed by Horizzon', 'Lead Quality_Worst'

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?
   'Tags_Lost to EINS', 'Tags_Closed by Horizzon', 'Lead Quality_Worst'

   'Tags_Lost to EINS', 'Tags_Closed by Horizzon' are obtained by encoding original categorical variable 'Tags'. 'Lead Quality_Worst' is obtained by encoding the categorical variable 'Lead Quality'.
   - Tags_Lost to EINS (Coefficient factor = 9.578632)
   - Tags_Closed by Horizzon (Coefficient factor = 8.555901)
   - Lead Quality_Worst (Coefficient factor =-3.943680)

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.

   build_model_cutoff(X_train[col], y_train, X_test[col], y_test, cutoff=0.1)

```
                                  Result of training data
   Prospect ID  Converted  Convert_Probability  Convert_predicted  Lead_Score
0         3009          0             0.289842                  1          29
1         1012          0             0.111387                  1          11
2         9226          0             0.001918                  0           0
3         4750          1             0.737087                  1          74
4         7987          1             0.993914                  1          99
------------------Result of test data-------------------
   Prospect ID  Converted  Convert_Probability  Convert_predicted  Lead_Score
0         3271          0             0.289842                  1          29
1         1490          1             0.929765                  1          93
2         7936          0             0.289842                  1          29
3         4216          1             0.998548                  1         100
4         3830          0             0.289842                  1          29
------------------Model Evaluation Metrics-------------------
Confusion Matrix :
 [[1221  513]
 [  44  945]]
Accuracy : 0.7954461990451708
Sensitivity : 0.9555106167846309
Specificity : 0.7041522491349481
Precision : 0.6481481481481481
```

	Prospect ID	Converted	Convert_Probability	Convert_predicted	Lead_Score
0	3271	0	0.289842	1	29
1	1490	1	0.929765	1	93
2	7936	0	0.289842	1	29
3	4216	1	0.998548	1	100
4	3830	0	0.289842	1	29
...	...	...	...	...	...
2718	850	0	0.070553	0	7
2719	2879	0	0.001642	0	0
2720	6501	1	0.989122	1	99
2721	7155	0	0.070553	0	7
2722	376	0	0.070553	0	7

2723 rows × 5 columns

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.

   During the quarter before the deadline, the company has very less time in its hand. Therefore, it is crucial to focus more on hot leads with the highest lead conversion rate. They ought to prioritize the leads and steer clear of pointless calls. Lead score can be used to determine priorities. It is possible to target leads with a lead score of greater than 80%.