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

Answer 1:

----------Feature Importance----------
const -1.965935
Total Time Spent on Website 1.679757
Lead Origin_Landing Page Submission -0.361672
Lead Origin_Lead Add Form 4.729656
Lead Source_Olark Chat 1.860251
Last Activity_Olark Chat Conversation -1.658297
Last Activity_Others -1.051297
Last Activity_Page Visited on Website -0.943932
What is your current occupation_Working Professional 1.449203
Tags_Will revert after reading the email 4.608920
Last Notable Activity_Others 0.845693
Last Notable Activity_SMS Sent 1.263945

Based on it

‘Lead Source_Olark Chat ’, : 1.860251
‘Lead Origin_Lead Add Form’, 4.729656
‘Tags_Will revert after reading the email ’ : 4.608920

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 2:

Same as first question as in this too all 3 are categorical/dummy variables

‘Lead Source_Olark Chat ’, : 1.860251
‘Lead Origin_Lead Add Form’, 4.729656
‘Tags_Will revert after reading the email ’ : 4.608920

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 3:
We will train the model using the lowest threshold= 0.1.
A lower threshold means more leads are classified as potential converts, ensuring a broader reach

Metric Score with optimal cut off =0.3

Accuracy: 0.8564
Sensitivity (Recall): 0.8650
Specificity: 0.8508
Precision: 0.7913

Results with Threshold 0.1:

Accuracy: 0.7583
Sensitivity (Recall): 0.9571
Specificity: 0.6283
Precision: 0.6274

Interpretation:

High Sensitivity: This indicates that the model is identifying almost all actual positives (leads likely to convert), which is crucial during the interns' period.

Moderate Specificity: While some non-leads are also being targeted, the high sensitivity ensures that the majority of potential leads are not missed.

Moderate Precision: While the precision is not very high, the aim here is to cast a wide net, which is acceptable during this aggressive conversion phase.

Given the high sensitivity, the interns will be busy, ensuring a higher chance of conversions.

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

We will train the model using the lowest threshold= 0.9
The sales team should make phone calls only to those leads that have a very high likelihood of conversion. This minimizes the number of calls made and ensures that the team focuses only on the most promising leads.

Metric Score with optimal cut off =0.3

Accuracy: 0.8564
Sensitivity (Recall): 0.8650
Specificity: 0.8508
Precision: 0.7913

Metric Score with optimal cut off =0.9
Accuracy : 0.8358585858585859
Sensitivity : 0.6085766423357665
Specificity : 0.9844868735083532
Precision : 0.9624819624819625

High Specificity (98.45%): Ensures almost all non-converting leads are correctly identified, reducing unnecessary calls.
High Precision (96.25%): Ensures that most of the contacted leads are likely to convert.
Trade-off: Lower Sensitivity (60.86%) means some potential leads might be missed, but this is acceptable when the goal is to reduce the number of calls.