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

Answer:

- Lead Source
- Total Time Spent on Website
- TotalVisits

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:

- Do Not Email
- Lead Source_Direct Traffic
- Lead Source_Organic Search

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:

This is achieved when the model correctly predicts the greatest number of people who are converted and also the number of people who are not converted i.e., the high sensitivity and high specificity, probability threshold of 0.3 will give better balance between sensitivity and specificity.

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:

To avoid useless calls the number of incorrect predictions should be less. This can be achieved by choosing the high specificity, if we choose the threshold 0.9 where the model will predict the highest number of correct predictions of not converted people.