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

**Answer:**

**‘Status\_reply after email’, Lead Origin\_Lead Add Form, Total Time Spent on Website are the top three variables in my model which contributes the highest towards lead conversion.**

1. 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:**

**‘Status\_reply after email’, Lead Origin\_Lead Add Form and Lead Source\_Welingak Website are the top 3 categorical variables which should be focused most in order to increase the probability of lead conversion**

1. 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:**

**Sensitivity is a measure which tells us the probability of positives being correctly identified out of the real positives. In this period we want to contact as many leads as possible without missing a slightest chance of conversion. So, here even if we call to a lead with low probability of conversion it is fine but we don’t want to miss out anyone.**

**Sensitivity = True positives/ (True positives + False negatives)**

**We decrease the probability as low as possible. According to our model 0.1 is a good cutoff to target maximum leads and ensure high sensitivity. But, it also depends on how many leads are the interns capable of targeting in total and then we can decide the cutoff accordingly**

1. 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 this case the company does not intend to spend time on making calls.**

**False Positive rate is a measure which tells the probability of a label being predicted as ‘yes’ when it is actually ‘no’ (label = 1 , labeling as a hot lead in this case ). We want False Positive rate to be as low as possible in other words we want specificity to be as high as possible**

**False Positive rate (FPR) = False positives/(False positives + True negatives)**

**Specificity = 1 - FPR**

**So, under such circumstances our specificity needs to be close to 1. This can be achieved by making the cut-off value of label assignment higher. Here, we may miss out some true potential leads but in this situation we can afford it**