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

Top 3 variables are :

1.Lead Source

2.Total Time Spent on Website

3.TotalVisits

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?

1.Lead Source\_Welingak Website

2.Lead Source\_Reference

3.Last Notable Activity\_Unreachable

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.

For making lead conversion more aggressive, all of the potential leads need to be identified as they have enough employees to make calls to leads. In this case the company is not concerned about the false positive rate and our aim is to increase Sensitivity/Recall. So, our objective here will be to reduce the number of false negative observations. That means we will have to decrease the cutoff for the predicted probability. This cutoff has to be confirmed with the help of business, considering achievable number of daily calls per employee.

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.

Here the aim is to reduce useless phone calls, which means they are not concerned about false negative values but are interested in high precision of the model. By increasing the cutoff of the model, the number of observations falsely identified as positive can be reduced. Thus, only those leads having high probability will be predicted as 1. Again this cut off will have to be confirmed with the business based on the number of phone calls they are planning to make per day.