

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

Solution: Lead Origin, Lead Source and Lead Profile are the 3 original features that have most positive effect on lead getting converted

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?

Solution: Lead Origin Lead Add Form, Lead Source_Welingak Website, Lead Profile_Lateral Student are the 3 variables which have the most positive effect on lead conversion And hence should be focused on.

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.

Solution: Since they have the man power to make the calls, the cutoff point can be chosen as the optimal point (which in this case is 0.35), which is the intersection of specificity and sensitivity and make all the calls to the individuals which have lead score greater than the cut off point. There can be some false positives but this is the trade off that can be made by the company given the scenario of sufficient man power.

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.

Solution : Given the man power is not high and company wants to reduce the number of phone calls, the strategy should be to choose a higher cut off value for lead score so that the number of false positives can be reduced. This might result in loss of some potential leads but it ensures that the probability of conversion is much higher since the leads with higher lead score are only selected.