

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

The top three variables which contribute most towards the probability of a lead getting converted are: Tags, Lead_Source, Last_Activity.

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?

The top three categorical/dummy variables that should be focused on the most in order to increase the probability of lead conversion are: Tags_Closed by Horizon, Tags_Lost to EINS, Tags_Will revert after reading the email

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.

For this scenario, ideally, not a single customer should be missed. Hence, it is important that the model identifies almost all the 'converted' customers correctly. It is fine if it incorrectly predicts some of the 'non-converted' customers as 'converted'. In that case, the worst that will happen is that the company will make phone calls to those customers who would not buy. In the model, the hot leads are defined as those that have more than 90 of Lead_score. In this case, X education might want to decrease the lead_score so they could aggressively increase lead conversion.

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.

X Education should focus on the precision rate – The probability that a predicted 'Yes' is actually a 'Yes'. Education can choose the cutoff probability threshold where the precision rate is maximized. X Education also can set a higher lead score as hot lead (maybe 93-95) so that they can reduce the number of pointless phone calls.