

Subjective Questions

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

Ans: In my opinion, the following are the top three variables:

The Total Time Spent on the website, Last activity and Lead source.

- The Total Time spent shows that those customers that spend more time learning and engaging more deeply with the content from the website have a higher probability of converting
- Last Activity for example opening the email that was sent to them, shows that they are again quite interested and have a higher chance of converting
- Lead Source like Google leads show that the customers are actively looking for a solution and are likely to take a step ahead and convert

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?

Ans: Lead Source, Lead Origin and Lead Quality are variables that should be focused on.

- Lead Source like in the previous question, Google Leads show a higher conversion probability. Thus the company should prioritize Google leads and focus more on paid campaigns and the like
- Lead origin such as landing page submission are more likely to convert as the customers that arrive show interest in the services
- Lead Quality is a given where 'High' quality have higher conversion rate. Hence the company should immediately prioritize with more immediate follow-ups.

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.

Ans: The following strategies can be applied:

- By default the threshold we use in logistic regression is 0.5, but by lowering it 0.3-0.4, the company can capture more leads that are more likely to convert which are classified as "Not Converted"
- Prioritizing higher engagement leads (ex: leads with high website time, opened emails, positive lead quality etc) should be always followed up and prioritized as they have a higher rate probability of conversion.

Subjective Questions

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.

Ans: If X Education reaches its quarterly target early and wishes to minimize unnecessary phone calls, the company should adopt the following strategy:

- Increase the probability threshold for classification as a converted lead. By setting a threshold closer to 0.7/0.8, only leads that have a very high chance of converting will be targeted for phone calls. This will ensure the focus is only on the most valuable leads and reduces time spent on less likely conversions.
- Focus on leads categorized as "Hot Leads" based on factors such as lead source, last activity (e.g., emails opened or clicked) and time spent on the website. These leads are more likely to convert, so efforts should be concentrated here.

This strategy will help the company minimize efforts while ensuring high-potential leads are still addressed.