

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

Top three variables:

- Total Time Spent on Website (coefficient: 5.762385)
- Lead Origin_Lead Add Form (coefficient: 4.024977)
- Avg_Time_Per_Visit (coefficient: -3.786653)

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?**

Top three categorical/dummy variables:

- Lead Origin_Lead Add Form (coefficient: 4.024977)
- Last Activity_Had a Phone Conversation (coefficient: 2.841200)
- Last Notable Activity_Unreachable (coefficient: 2.738682)

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.**

To increase the potential leads, the optimal way is decreasing the threshold.

However, the strategy should go more than that with the combinations:

First, decreasing the threshold. Decrease the threshold from best model from 0.42 to 0.3 or even 0.2. More potential converters get-in, increase the recall but also maybe lower the precision which increases the false negative

Second, outreach the expanded database. As the threshold has decreased, more and more customers will enter in the database, we can use the additional people (10 interns) to approach the customers as much as possible makes it feasible to contact nearly everyone flagged by the model.

Third, segmenting the customer with the probabilities to buy. For the core sales team members, leads with the highest probability will send them for approaching, while interns would approach the lower probability which can be the case of False Positive. By applying model best-to-best, no customers has been left behind also the lower probabilities can have the follow-up calls later on as Last Notable Activity_Unreachable also being one of the good dummy variables impact to model

Last one, adjust the model or transfer best customer to some interns have been performed better than the core team.

- 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.**

The strategy will be reversed with question 3. We need to increase the threshold to tightening the criteria of leads classified as "hot". The strategy will follow the steps below:

First, increasing the threshold. Increase the threshold from best model from 0.42 to 0.6 or even 0.8. The sales team will only approach the clients being extremely converted

Second, limit all the clients having lower probabilities to be called. It would lead to the rate of useless phone calls decreasing.

Third, monitor the conversion rate. By increasing the threshold, the conversion rate is expected to be higher as the outreached clients now have very high probability. Some quality checks and model modification would be needed if the conversion rate is not increased as expected.