

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

Answer:

The top three variables in the model which contribute most towards the lead getting converted include:

- **Total visits** – The more visits would mean the lead is curious to know more
- **Total Time Spent on Website** – More time spent trying to understand the courses would make it easy to persuade the lead and converting it
- **Lead source: Google Search** – Search through google has higher conversion rate

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?

Answer:

The top three categorical variables in the model which should be focused are:

- **Lead source API** – Clark Chat. Connect with the lead is a primary way to explain about the courses and it is the best opportunity for X Education to convert the lead into a potential client.
- **Lead source API** – Organic Search – Organic search is one of the ways to know the interests of the lead and these leads would have a higher potential to convert based on their interest
- **Lead source Landing Page Submission** – Direct Traffic – Direct traffic is more significant to understand the leads interest in pursuing the course from X Education and it has a greater potential to convert the lead

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.

Answer:

During this period they can reduce the cut off for lead identification as the aim is to identify all the potential leads and at the same time, resource is not a challenge. This shall help to increase the Recall which measures the proportion of actual positives (converted leads) that the model correctly identifies. Also, for aggressive campaign, it is important to identify as many potential leads as possible, even if it means some false positives (non-converted leads are incorrectly classified as potential).

Hence, focusing on recall ensures that you do not miss out on leads who could potentially convert.

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.

Answer:

In this case as the company wants to be selective in making the call, they can increase the threshold for lead prediction cutoff. This shall increase the **Precision** of the model which is correctly predicted positive leads (True Positives, TP) out of all predicted positive leads (TP + FP). This shall help focusing on "hot" leads or leads with high probability of conversion. This shall thus help in minimizing useless phone calls.