

LEAD SCORING CASE STUDY

ASSIGNMENT SUBJECTIVE QUESTIONS

Prasasti Choudhury, Satadhriti Chakrabarty
DS C17 Group 2 February 2020

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 contributing most towards the possibility of a lead getting converted are **Tags**, **Lead Source** and **Lead Origin**.

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 contributing most towards the possibility of a lead getting converted are **Tags_Closed by Horizzon**, **Tags_Will revert after reading the email** and **Tags_Not Mentioned**.

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.
 - The interns should also look at the Lead Score. Based on our model, any customer with a Conversion probability more than 30% will be predicted as 1. But a person with a conversion probability of, say around 50%, will have a lesser chance of getting converted. However, if the company is decided on the call that they will call all customers predicted as 1 by the model, they should focus on the key driving factors of a lead getting converted from the model. A **combination of the model results** where we can see the coefficients in the logistic regression model and the **EDA results** will act as a good guide to do some prerequisites before making a phone call to someone. For example, in case of

Lead origin, the focus should be more on the **Lead Add Form** as it has a high conversion rate. In case of **Tags**, customers who want to '**Revert after reading the email**' should be targeted aggressively with additional discounts and offers. People with **Last Notable Activity** as '**SMS sent**' have higher chances of getting converted. So, a **mix of both the EDA results and the final model coefficients** will act as a good starting point to ensure that more and more leads get converted.

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 ideal strategy should be to look at the **Lead_Score** generated by the model before calling. A cutoff point should be determined (**ideally around 80**) and people having LeadScores above that point should be the only targets. These are **Hot Leads** and have high chances of getting converted. This **reduces** the **effort** of the sales team to a great extent.