# Lead Scoring Case Study

**Question: 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 in the Logistic Regression are

1. Tags
2. What is your current occupation
3. Last Activity

- Tags – Explains about the Tags assigned to customers indicating the current status of the lead.
- What is the current occupation – Explains about the Lead's occupation.
- Last Activity - Explains about the last activity of the customer.

**Question: 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 in the model that should be focused are -

1. Tags_Lost to EINS
2. What is your current occupation_Working Professional
3. Last Activity_SMS sent

- Tags_Lost to EINS: Leads who 'Lost to EINS category' is highly corelated to the lead's conversion.
- What is your current occupation_Working Professional: The leads who are 'working' are more likely to join the X Educational Organization.
- Last Activity_SMS sent – The Leads whose last activity is 'SMS sent' are more likely to convert

**Question – 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 strategy to be followed is to decrease the threshold value or the cut off value so that there will be high sensitivity. That means almost all the leads who are likely to be converted are identified.

In other words, if cut off value decreases there will be less accuracy, less specificity but more sensitivity.

High sensitivity which means that almost all leads who are likely to convert are identified and can be followed up by the new interns making the conversion rate high.

**Question – 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.**

High Specifity means the model will correctly identify almost all leads who are not likely to convert. Leads which has chances of low conversion rates are not identified as the hot leads.

As X Education has already reached its target for a quarter and not aiming in making calls unless necessary, we can choose high threshold value for the conversion probability making the false positive rate less. (1-specificity)

High specificity means the leads who are on the edge of the conversion are not selected. Making the agents to make only necessary calls and can focus on new work.