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

Ans : Top three most contributing variables in our model are :

a. ‘Last notable activity\_SMS sent’,

b. ‘Lead quality\_not sure’

c. ‘Tags\_will revert after reading the email’

‘Last notable activity\_SMS sent’ has a correlation value of 0.36 with ‘Converted’

‘Lead quality\_not sure’ has a negative correlation value of -0.44 with ‘Converted’

‘Tags\_will revert after reading the email’ has a correlation value of 0.35 with ‘Converted’

1. 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 : Top 3 categorical/dummy variables in the model to be focused in order to increase probability of lead conversion are :

* 1. Tags
  2. Last Notable Activity
  3. Lead Quality

Tags(Current Status): Tags with "will revert after reading the email have high leads as well as high conversion rates."​

Last Notable activity: More number of leads have Modified as Last notable activity however with very low conversion rate, leads with 'SMS Sent' have high Conversion rate.

Lead Quality : Lead Quality with ‘Not Sure’ has a very high negative correlation with converted field.

1. 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 : A good strategy to employ at this stage is to **select cut-off value low (0.3)**, so that most the cases will be predicted as ‘Converted’. The cases with even low probability will be predicted as converted too.

We should increase the sensitivity.

High sensitivity implies that our model will correctly identify almost all customers who are likely to convert. It will do that by over-estimating the converted likelihood, i.e. it will misclassify some non-converters as converters, but that is the trade-off we need to choose rather than the opposite case.

And by selecting the cut-off value low, we can increase the sensitivity.

1. 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 : A good strategy to employ at this stage is to **select cut-off value high (0.5)**, so that less cases will be predicted as ‘Converted’. The cases with high probability will be predicted as converted only.

We are decreasing the sensitivity and increasing the specificity.

High specificity implies that our model will correctly identify almost all customers who are not likely to convert. It will do that by over-estimating the non-converted likelihood, i.e. it will misclassify some converters as non-converters, but that is the trade-off we need to choose rather than the opposite case.

And by selecting the cut-off value high, we can increase the specificity.