1. Which are the top three variables in your model which contribute most towards the probability of a lead getting converted?
   1. The three main features are lead origin, tags and current\_occupation
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?
   1. The three categorical features are lead\_origin\_other, tags\_ringing, current\_occupation\_working \_professional
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.
   1. Focus on a model with more recall. So the model will reduce the false negatives and we’ll have more potential leads. Note that this will also have more TN or more number of students who won’t convert.
   2. By checking the model parameters, we can identify why a particular candidate was predicted as Positive. We can ask the interns to emphasise on those attributes/details while making the calls. For example if the student is a working professional, the intern can talk about the benefits a particular course provides for working professional.
   3. Improve the model by adding more data on a daily basis and then using those models for subsequent calls.
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.
   1. We can create a model where the FP is very low. That means the model will rarely predict an FP or in other words the model will only predict the student to be converted if it has high confidence. This comes at the cost of having a higher FN.