Lead Scoring Assignment – Pappu Kapgate & Shivram J

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

   Top three variable which contributed most towards the probability are:

   - Total Time Spent on Website
   - What is your current occupation?
   - 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?

   Top 3 category/dummy variables are:

   - Lead Origin_lead add form
   - What is your current occupation_working professional
   - Lead Source_welingak website

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.

   A good strategy in this case would be to lower the threshold [cut off] for conversion probability so that more leads can be pursued by the interns. This will broaden their reach to make more phone calls. In the final output of the model where each lead is represented with a lead score, the interns can start looking at leads with lower scores to increase lead conversion rate.

   Some additional factors to look at is Total Time Spent on Website, Lead source and profession when looking at filtering leads.

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.

   In this case we make the model to be more sensitive so that only higher probability leads are pursed, and unnecessary phone calls are not made. We can again leverage lead

score to identify higher probable hot leads from the pool of lead which have been already classified as a good lead by the model.