Subjective Questions

Contributors: Vidhi Surana & Shubham (Batch DS C48)

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

Ans: The top three variables in our model which contribute most towards the probability of a lead getting converted are:

a. Lead Source_Welingak Website
   Focusing more on how to advertise more on Welingak can attract more leads.
b. Lead Source_Reference
   We can provide discount/reward to people for providing references.
c. What is your current occupation_Working Professional
   We should engage working professional through communication channels based on their engagement impact.

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?**

Ans: 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 are:

a. Lead Source_Welingak Website
b. Lead Source_Reference
c. What is your 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.***

Ans: So, Sensitivity is a measure of predicting yes when its actually yes. It is defined as

$$\text{Sensitivity} = \frac{\text{Number of Actual Yeses correctly predicted}}{\text{Total Number of Actual Yeses}}$$

As the optimal threshold decreases, sensitivity increases and specificity decreases and vice versa.
To ensure that almost all the potential leads are converted we need to ensure the sensitivity is more and for that we will have to adjust the threshold. If we experiment and *decrease the optimal threshold,* sensitivity keeps on increasing. In turn increasing the number of hot leads identified. Since there are more people in the sales team for 2 months, they will have the time to call more number of leads. Increase in sensitivity will increase the number of hot leads identified. So, that the sales team can call them aggressively. It might happen that because of high sensitivity some leads are non-converting will also be contacted but that's something which wont bother as there is lot of staff for 2 months.

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***.

Ans: Since Company has already reached its target for a quarter and wants to save the time that will otherwise be wasted by contacting leads that will not convert, it will be better if they contact only those leads which are really going to convert. For this, we should make sure that the specificity of the model is high,

$$\text{Specificity} = \frac{\text{Number of actual Nos correctly predicted}}{\text{Total number of actual Nos}}$$

This can happen only when we *increase the optimal threshold*, which in turn decreases the sensitivity and increases the specificity. This will make sure that people that will not convert will not get selected at all.