Lead Scoring Case Study

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

   Answer:

   We have built the Logistics Regression model with below features and in this feature top 3 variables are below:

   I. **Tags_Lost to EINS**
   II. **Tags_Closed by Horizons**
   III. **Tags_Will revert after reading the email**

```
feature_importance.sort_values(ascending=False)
```

```
Tags_Lost to EINS                                   100.00
Tags_Closed by Horizzon                              86.68
Tags_Will revert after reading the email             43.22
Tags_Busy                                            41.44
Lead Source_Welingak Website                         37.19
Last Notable Activity_SMS Sent                       29.86
What is your current occupation_Working Professional 14.60
Lead Origin_Lead Add Form                            11.87
Do Not Email                                        -14.18
Tags_Ringing                                        -18.46
Tags_switched off                                   -24.93
Lead Quality_Not Sure                               -36.40
Lead Quality_Worst                                  -40.99
dtype: float64
```

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?**

   Answer:

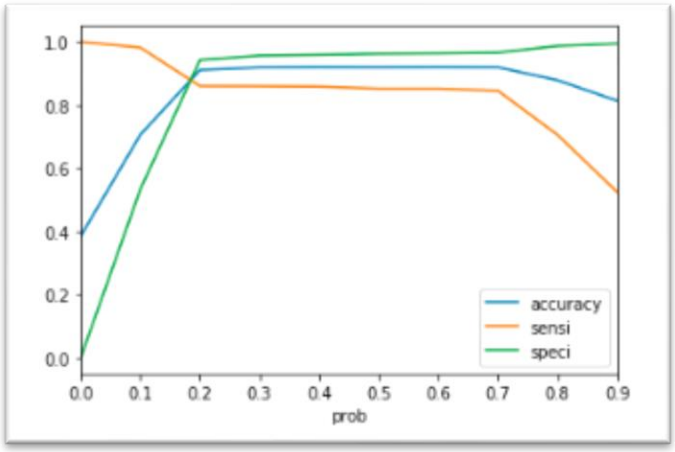
   As per the above list the top three categorical/dummy variables in the model which should be focused the most on in order to increase the probability of lead conversion are:

   I. **Tags_Lost to EINS**
   II. **Tags_Closed by Horizzon**
   III. **Tags_Will revert after reading the email**

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.**

Answer:

As per the model we got the below Cutoff points.
Note: Optimal cutoff probability is that prob where we get balanced sensitivity and specificity



They need to make the lead conversion more aggressive. For that we need to understand the Sensitivity and Specificity from this graph.

Sensitivity is defined as ratio of total number of actual Converted correctly predicted to the total number of actual Converted.

Specificity is defined as the ratio of total number of actual non-Converted correctly predicted to the total number of actual non-Converted.

So, these two are reciprocal, if one is increase other will decrease.

As we have focus to convert the lead more aggressive so we need to make sure the Sensitivity should be high means Cutoff should be low. High sensitivity implies that our model will correctly identify almost all leads who are likely to convert.

This will give them a greater number of leads and also, they have enough staff to make more calls and this will help **them make the lead conversion more aggressive.**

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.**

   Answer:

   The company's aim is to not make phone calls unless it's extremely necessary as the company reaches its target for a quarter before the deadline.

   High Specificity implies that our model will correctly identify almost all leads who are not likely to convert.

   Here, also the same concept but this they need accurate leads. Here we will make the cutoff high so that we have less number projected leads.

   So Here the precision would be much high as we are contacting the hot leads which has high chances of conversion and will achieve the aim to avoid useless phone calls.

Team:

- Rohit Keshari
- Rahul Choudhary

---------End of Subjective Questions---------