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

   ANS : Our final model has about 15 columns and most of these are from the 3 categories mentioned below.
   1) "What is your current occupation"
   2) "Last Notable Activity"
   3) "Lead Origin"

   Hence these are the primary columns which contribute towards the lead getting converted. To be very specific about the variables, below are the three variables which contribute the most towards the lead conversion probability
   1) "Occupation_Working Professional"
   2) "Lst_Not_Act_Had a Phone Conversation"
   3) "LeadOrigin_Lead Add Form"

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 : Based on the model summary, we should be focusing on the columns with the high values of coefficient as these mainly contribute towards the lead getting converted. Hence the top 3 categorical/dummy variables in order of priority are:
   - ✓ "Occupation_Working Professional"
   - ✓ "Lst_Not_Act_Had a Phone Conversation"
   - ✓ "LeadOrigin_Lead Add Form"

   The above dummy variables are derived from the following categorical columns:
   1. "What is your current occupation" : answered as "Working Professional"
   2. "Last Notable Activity" : registered as "Had a Phone Conversation"
   3. "Lead Origin: : identified as "Lead Add Form"

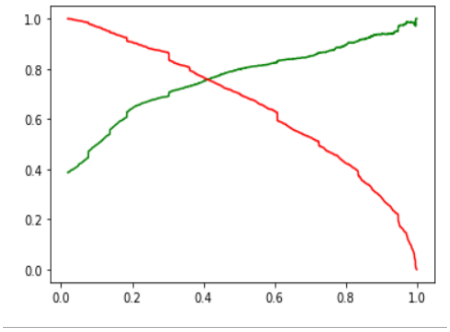
3. **X Education has a period of 2 months every year during which they hire some interns. The sales team, in particular, has around 10 interns allotted to them. So during this phase, they wish to make the lead conversion more aggressive. So they want almost all of the potential leads (i.e. the customers 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 : Adding 10 more interns to the sales team results in increased capacity of the sales team to make more no of calls. A good strategy to increase the number of potential leads that sales team will reach out to is as follows: The present lead score and prediction values have been derived by using an optimal cutoff probability of 0.35 which gives accuracy and sensitivity values at around 81%. We can decrease the cut-off value while doing the predictions from 0.35 to a lower value as depicted in the table. This will have an impact on the precision, it will increase the number of potential projected leads

on our dataset thereby providing more number of customers whom the calls can be made using the interns.
While reducing the cut-off value we should review the other metrics like accuracy and specificity etc. just to make sure that they do not fall below the acceptable limits.

Precision Versus Recall and other parameters

	prob	accuracy	sensi	speci	preci
0.00	0.00	0.381262	1.000000	0.000000	0.381262
0.05	0.05	0.473871	0.992295	0.154423	0.419654
0.10	0.10	0.615028	0.967153	0.398051	0.497497
0.15	0.15	0.710266	0.940389	0.568466	0.573159
0.20	0.20	0.772109	0.906732	0.689155	0.642529
0.25	0.25	0.792053	0.881184	0.737131	0.673798
0.30	0.30	0.801330	0.865775	0.761619	0.691162
0.35	0.35	0.811843	0.811030	0.812344	0.727008
0.40	0.40	0.814162	0.767234	0.843078	0.750794
0.45	0.45	0.818182	0.732360	0.871064	0.777778
0.50	0.50	0.818955	0.701946	0.891054	0.798800
0.55	0.55	0.814162	0.667883	0.904298	0.811330
0.60	0.60	0.808442	0.631792	0.917291	0.824775
0.65	0.65	0.791435	0.560016	0.934033	0.839514
0.70	0.70	0.784941	0.527575	0.943528	0.851997
0.75	0.75	0.771336	0.470803	0.956522	0.869663
0.80	0.80	0.761441	0.423763	0.969515	0.895458
0.85	0.85	0.741033	0.352393	0.980510	0.917635
0.90	0.90	0.721243	0.287916	0.988256	0.937913



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 : The requirement is to minimize the rate of useless phone calls. It also means that we want to only focus on making calls to those candidates who have a very high probability of conversion. A good strategy to help in such situation is to increase the cut-off value from the optimal value of 0.35 to a higher value as per the table below. This will increase the precision (accuracy of the detected positives) for the projected values thereby ensuring a higher conversion rate for every phone call. As we increase the cut-off, it may further impact the accuracy and the False Negative ratio, So a decision should be taken based on the tolerance for these values.

Precision Versus Recall and other parameters

	prob	accuracy	sensi	speci	preci
0.00	0.00	0.381262	1.000000	0.000000	0.381262
0.05	0.05	0.473871	0.992295	0.154423	0.419654
0.10	0.10	0.615028	0.967153	0.398051	0.497497
0.15	0.15	0.710266	0.940389	0.568466	0.573159
0.20	0.20	0.772109	0.906732	0.689155	0.642529
0.25	0.25	0.792053	0.881184	0.737131	0.673798
0.30	0.30	0.801330	0.865775	0.761619	0.691162
0.35	0.35	0.811843	0.811030	0.812344	0.727008
0.40	0.40	0.814162	0.767234	0.843078	0.750794
0.45	0.45	0.818182	0.732360	0.871064	0.777778
0.50	0.50	0.818955	0.701946	0.891054	0.798800
0.55	0.55	0.814162	0.667883	0.904298	0.811330
0.60	0.60	0.808442	0.631792	0.917291	0.824775
0.65	0.65	0.791435	0.560016	0.934033	0.839514
0.70	0.70	0.784941	0.527575	0.943528	0.851997
0.75	0.75	0.771336	0.470803	0.956522	0.869663
0.80	0.80	0.761441	0.423763	0.969515	0.895458
0.85	0.85	0.741033	0.352393	0.980510	0.917635
0.90	0.90	0.721243	0.287916	0.988256	0.937913