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

Answer: Following are the top three variables that contribute most toward the probability of a lead getting converted:

- Lead Source
- What is your current occupation
- Last Activity.

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: Following are the top 3 categorical/ dummy variables that should be focused the most on in order to increase the probability of lead conversion:

- Lead Source_Welingak Website
- Lead Source_Reference
- 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.**

Answer:

Here, the concept of sensitivity is required, ***Sensitivity* = *T*P/ (*T*P + *F*N) :**

Sensitivity can be defined as the number of actual conversions predicted correctly out of total number of actual conversions.

As we know, sensitivity decreases as the threshold increases. In the given situation, we'll need a **high sensitivity** because high sensitivity will mean that our model will correctly predict almost all leads who are likely to convert. To make the lead conversion more aggressive by making phone calls to as many potential leads as possible, it is a good strategy to go for high sensitivity. To achieve high sensitivity, we need to **choose a low threshold value.** Below Table of pic for reference.

	prob	accuracy	sensi	speci
0.0	0.0	0.385136	1.000000	0.000000
0.1	0.1	0.586049	0.973426	0.343406
0.2	0.2	0.748386	0.916599	0.643022
0.3	0.3	0.801449	0.849959	0.771063
0.4	0.4	0.816564	0.772690	0.844046
0.5	0.5	0.816879	0.706051	0.886300
0.6	0.6	0.797040	0.600572	0.920102
0.7	0.7	0.786963	0.531889	0.946735
0.8	0.8	0.761297	0.424775	0.972087
0.9	0.9	0.720831	0.294767	0.987708

With respect to our model, The company **should make calls** to the leads by considering the below listed features, as these are more likely to get converted.

- Leads coming from the **lead sources: Welingak Websites** and **Reference**.
- Leads who are the **working professionals**.
- Leads whose **last activity** was **SMS Sent**.
- Leads who spent **more time on the websites**.
- Leads coming from the **lead sources Olark Chat**.

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:

Here, the concept of specificity is required, ***Specificity* = *T*N/ (*T*N + *F*P)**

Specificity can be defined as the number of actual non-conversions predicted correctly out of the total number of actual non-conversions.

As we know that the specificity increases as the threshold increases. In the given situation, we'll need a **high specificity** because high specificity will mean that our model will correctly predict almost all leads who are not likely to convert.

If a company doesn't want to make phone calls unless it's extremely necessary, it is a good strategy to go for high specificity. It will ensure that the phone calls are only made to customers who have a very high probability of conversion. To achieve high specificity, we need to **choose a high threshold value**. Below Table of pic for reference.

	prob	accuracy	sensi	speci
0.0	0.0	0.385136	1.000000	0.000000
0.1	0.1	0.586049	0.973426	0.343406
0.2	0.2	0.748386	0.916599	0.643022
0.3	0.3	0.801449	0.849959	0.771063
0.4	0.4	0.816564	0.772690	0.844046
0.5	0.5	0.816879	0.706051	0.886300
0.6	0.6	0.797040	0.600572	0.920102
0.7	0.7	0.786963	0.531889	0.946735
0.8	0.8	0.761297	0.424775	0.972087
0.9	0.9	0.720831	0.294767	0.987708

With respect to our model, During the quarter before the deadline, the company has very less time in its hand. So it is of utmost importance that it concentrates more on **hot leads that have the highest lead conversion rate**. They should avoid useless calls and prioritize the leads. Prioritizing can be done on the basis of lead score. Leads that have more than 80% of lead score can be targeted.