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

Answer: Based on the coefficients for the last model trained in the iteration process below are the top 3 variables which contributes most towards lead conversion

Total Time Spent on Website – Leads spending more time on the website have higher chances of conversion.

Lead Origin_**Lead Add Form** – Leads originating from the source Add Form have higher chances of conversion.

What is your current occupation_**Working Professional** – Leads who are working professionals have higher chances of conversion.

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: Based on the coefficients for the last models trained in the iteration process, below are the top 3 categorical variables which increases the probability of lead conversion

Lead Origin_**Lead Add Form** – Leads originating from Add Form lead origin.

What is your current occupation_**Working Professional** – Leads who are working professional.

Lead Source_**Welingak Website** – Source of the lead is from 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.

Answer: As part of this project, we have come out with a good logistic regression model who has a good accuracy above 80% and a good recall also above 80%. Recommendation for Xeducation is to create a data pipeline to get the lead data and based on the leads generated, run the data against the Model built and get the predictions. The model will give a probability between 0 to 1 for all the leads and the leads who have a probability higher than 0.36, Xeducation should focus all of its sales team and sales interns make phone calls to the leads whose conversion probability is above 0.36. Out of the data ingested to the model, based on the training and test data we could almost see more than 80% of the leads

coming as converted which is a good sign. Based on the actual data if Xeducation feels the conversion numbers are less they can focus on reducing the probability cut off to 0.3 or 0.25. Also the sales and sales interns should focus on the lead aspects like more time spent on the website, Add form leads and leads getting generated from the Welingak Website and Olark chat and focus calling these leads before any other 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.

Answer: As part of this project, we have come out with a good logistic regression model who has a good accuracy above 80% and a good recall also above 80%. Recommendation for Xeducation is to create a data pipeline to get the lead data and based on the leads generated, run the data against the Model built and get the predictions. The model will give a probability between 0 to 1 for all the leads and the leads who have a probability higher than 0.36, Xeducation should focus all of its sales team and sales interns make phone calls to the leads whose conversion probability is above 0.36. If the sales team wants to focus on other activities they need not call the leads whose conversion probability is less than 0.36 and only focus on those leads whose conversion probability is above 0.36. As mentioned above, time spent on the website, Add Form lead origin, lead source as Welingak website and Olark chat specific leads should only be focused after the targets have been met by the sales team. This will ensure the lead generation records are optimized and the sales team can focus on other activities.