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

Answer:
➢ As per the final model there are 14 fields that significantly contribute towards the probability of a lead getting converted
➢ Out of these 14, following are the three top variables:

- Wherever the "Lead Origin" is "Others" which means either of (Lead Add Form, Lead Import, Quick Add Form) then it increases the "log of odds" of getting converted by 3.6 times
- If the lead is a "Working professional" it increases the "log of odds" of getting converted by 2.7 times
- If the last activity performed by the customer is "SMS Sent" then it increases the "log of odds" of getting converted by 1.26 times

Do Not Email	-1.1763
TotalVisits	0.1594
Total Time Spent on Website	1.0311
Lead Origin_Landing Page Submission	-0.4113
Lead Origin_Others	3.6057
Last Notable Activity_Modified	-0.8604
What is your current occupation_Student	0.5524
What is your current occupation_Working Professional	2.7175
Last Activity_Olark Chat Conversation	-0.9595
Last Activity_Page Visited on Website	-0.3573
Last Activity_SMS Sent	1.2559
Lead Source_Google	0.1831
Lead Source_Olark Chat	1.1586
Specialization_Others	-0.4002

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:

Focus should be on:
➢ What was the "Lead Origin" - in particular if it is
- Lead Add Form
- Lead Import
- Quick Add Form)
➢ What is the "Current Occupation" – in particular if it is
- Working Professional
➢ What was the "Last activity" done by the customer – in particular if it is
- SMS Sent

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:

➢ For the two months, where additional sales interns have been allotted, the lead conversion has to be more aggressive

➢ This means the company is ok if some "False positives" (i.e. some customer that have been predicted by the model as Hot leads – whereas they are not) are contacted, but even if there is some chance of a customer to be a hot lead – company would like them to be contacted

➢ This means we want a higher "Sensitivity" even if specificity goes down (which means False Positives would also increase)

➢ So, we can change the cut-off from 0.3 to 0.1

➢ This would take care of all the true positives, although would increase the false postives also – but at least greater sales capacity could be deployed to cover more potential hot leads

➢ In addition, some qualititative feedback can be taken care of. As per EDA it was found that Lead spending more time on the website are more likely to be converted. Hence website should be made more engaging to increase conversions. Sales interns could work with the technical teams to get these improvements planned.

➢ Leads who are seen visiting the website frequently can be considered as hot leads and should be targeted for conversion.

➢ Working professionals are hot leads, and hence can be targeted for easier conversion.
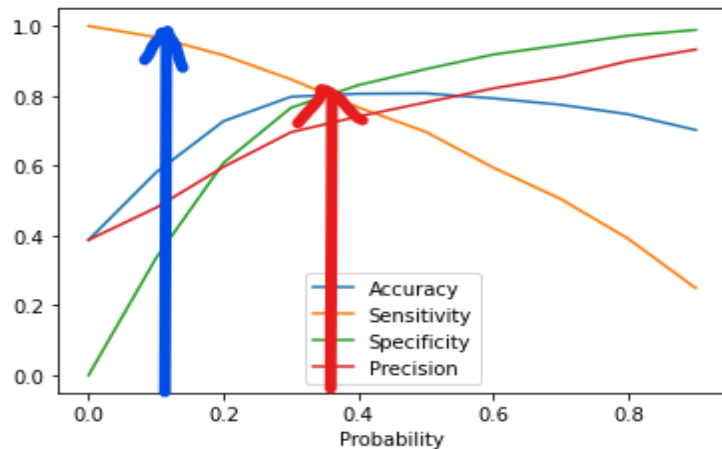
4. Similarly, at times, the company reaches its target for a quarter before the deadline. During this time, the company wants the sales team to focus on some new work as well. So, during this time, the company's aim is to not make phone calls unless it's extremely necessary, i.e. they want to minimize the rate of useless phone calls. Suggest a strategy they should employ at this stage.

Answer:

➢ On the contrary to Question 3, this is where the company wants to be selective i.e. have limited sales capacity be deployed to contact the customers

➢ Hence, we wanted to be really specific about the ML recommendations, in case a customer is expected not to be converted – we should be accurate enough to tell that, hence "Specificity" becomes important (which is true negatives/true negatives + false positives)

➢ Even if some customers who have a little potential to convert are missed upon, that is also fine (as the company wants to make phone calls only when extremely necessary)

➢ Hence, in this case the predictions should be done considering a at a higher cut-off.

➢ As per the attached graph – it could be 0.7 (or even higher based on company's choice) - this would ensure true negatives are captured well

➢ Email, Message, Chatbot can be other alternatives for call. All are automated solutions and hence need less manual effort.

➢ Calls should be targeted only for hot leads: customers with high Lead Score because they are more likely to buy the course.

➢ Providing enough convincing information on the website so that customers need not be convinced via call.