Q1. Which are the top three variables in your model which contribute most towards the probability of a lead getting converted?

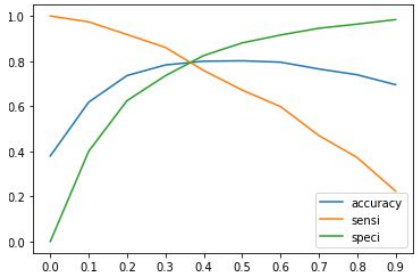
A1. Top 3 variables in the model are – Lead Source_WelingakWebsite, LeadSource_Reference, LastActivity_SMS Sent

Q2. 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?

A2. Top 3 variables that should be focused on to increase probability of lead conversion are – Lead Source_WelingakWebsite, LeadSource_Reference, LastActivity_SMS Sent.

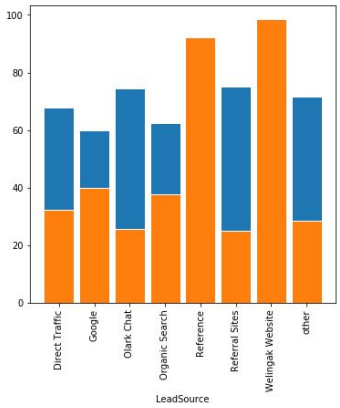
Q3. X Education has a period of 2 months every year during which they hire some interns. The sales team has around 10 interns allotted to them. So, during this phase, they wish to make the lead conversion more aggressive. So, they want almost all of the potential leads (i.e. the customers who have been 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.

A3. Technically, for this case we need to maximize the positives predicted by the model i.e. TPR or sensitivity or recall should be maximum. We can re-adjust model to use a different cut-off probably 0.1 or 0.2. This will help in increasing the scope of leads to be followed.



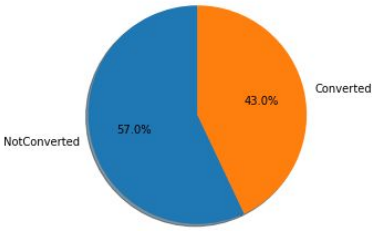
However, we have following observations from data.

1. People coming from references and Welingak Website have highest conversion rate. However, we still loose about 8% of this traffic. Experienced marketing team can focus on this group.

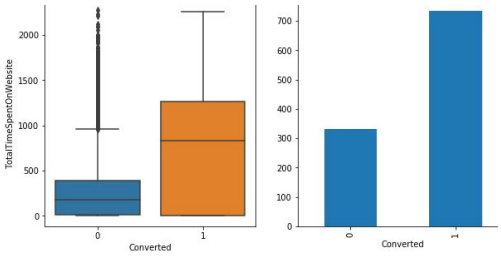


2. People coming through google search are next in the list. (40% conversion rate). These people should be targeted with more aggressive manner as there is a probability of loosing these people to competitors. These people are looking for something relevant and should be approached with solutions to their queries.
3. People who have been reached via SMS have high probability of conversion. However, we are still missing out on 30% of the audience. Interns can target these people via calling.
4. Working professionals are more inclined towards online courses, as they need flexible learning schedules and career transitions. We are losing around 8% of the traffic here. Experience marketing professionals can try to fix the leaks here or understand the reason for non-conversion.

5. Similarly, we have a conversion rate of around 43% for unemployed people. However, the number of leads is highest in this category. This is huge opportunity size. The only challenge with these people could be costing and financing. If the company can provide flexible payment options, placement assistance etc. this can increase conversion rate a lot and end results will lead to publicity by word of mouth and we have already seen that people coming through reference have high probability of conversion. Interns can be deployed here for repetitive follow-ups.



6. Mean time spent on the website by the leads that were converted was 700 units. Among the unconverted leads 15% of the people spent above average time on the website. These can be a good target audience for interns to work on and get them converted. Feedback from these people will also help in improving website/content.



Q4. 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 is extremely necessary, i.e. they want to minimize the rate of useless phone calls. Suggest a strategy they should employ at this stage.

A4. Technically, we would want to reduce the rate of false positives predicted by the model i.e. minimize the number of non-conversions being predicted as conversions. In other words, we must increase our precision for predicting the actual conversions. This can be done by increasing the cut-off for the model.

We have assigned lead score to every lead record. Team should make calls only to leads with score>0.9. These are around 10% of the data.

Another way to go about it could be filtering the data by features that we have observed to be key in determining the conversion of lead. For example, filtering the data on Lead source and last notable activity, leads sourced from - Welingak Website and Reference, and Last notable activity - SMS sent are favourable leads.