

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

- The three variables from the model that contribute towards the probability of a lead getting converted are:
- Total Time Spent on Website
- Page Views Per Visit
- 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?

- Last Notable Activity_Olark Chat Conversation
- Tags_switched off
- Last Activity_SMS Sent

	Features	VIF
12	Last Notable Activity_Olark Chat Conversation	2.67
9	Tags_switched off	2.03
1	Last Activity_SMS Sent	1.52
3	Tags_Busy	1.38
10	Lead Quality_Not Sure	1.35
0	Do Not Email	1.31
11	Lead Quality_Worst	1.28
8	Tags_Will revert after reading the email	1.25
2	Specialization_IT Projects Management	1.15
6	Tags_Not doing further education	1.10
5	Tags_Lost to EINS	1.09
7	Tags_Ringing	1.07
4	Tags_Interested in full time MBA	1.00

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.

It is understood from the analysis, columns that have higher “Total Time Spent on Website”, “Page Views Per Visit” indicates that the leads are interested in the service and it better to employ the people basis of their Yes response.

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.

Since the specificity and sensitivity are good to consider from the analysis, we can consider outputs from “Tags_Will revert after reading the email”, which means people will be interested in a course, though there are not enough calls been made.

```
# Let us calculate specificity  
TN / float(TN+FP)
```

```
0.8589420654911839
```

```
# Let's see the sensitivity of our logistic regression model  
TP / float(TP+FN)
```

```
0.9314814814814815
```