

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

**Answer**

Below is the final model used for the prediction:

Generalized Linear Model Regression Results						
=====						
Dep. Variable:	Converted	No. Observations:	3463			
Model:	GLM	Df Residuals:	3446			
Model Family:	Binomial	Df Model:	16			
Link Function:	logit	Scale:	1.0000			
Method:	IRLS	Log-Likelihood:	-1388.9			
Date:	Sun, 17 Nov 2019	Deviance:	2777.9			
Time:	15:26:57	Pearson chi2:	3.64e+03			
No. Iterations:	6	Covariance Type:	nonrobust			
=====						
	coef	std err	z	P> z	[0.025	0.975]
-----						
const	-3.1156	0.207	-15.062	0.000	-3.521	-2.710
Do Not Email	-1.6588	0.319	-5.204	0.000	-2.284	-1.034
Total Time Spent on Website	1.0974	0.051	21.391	0.000	0.997	1.198
Asymmetrique Activity Score	1.8648	0.113	16.545	0.000	1.644	2.086
Lead Origin_Lead Add Form	3.1021	0.263	11.790	0.000	2.586	3.618
Lead Source_Welingak Website	2.0074	0.858	2.339	0.019	0.325	3.689
Last Activity_Converted to Lead	-1.1579	0.237	-4.885	0.000	-1.622	-0.693
Last Activity_Email Bounced	-1.5930	0.538	-2.961	0.003	-2.648	-0.538
Last Activity_Had a Phone Conversation	2.5781	0.891	2.895	0.004	0.833	4.324
Last Activity_Olark Chat Conversation	-1.7234	0.198	-8.684	0.000	-2.112	-1.334
Last Activity_Page Visited on Website	-1.0603	0.207	-5.132	0.000	-1.465	-0.655
Last Activity_Unsubscribed	1.8221	0.693	2.629	0.009	0.464	3.180
Asymmetrique Activity Index_Low	4.3067	0.452	9.534	0.000	3.421	5.192
Asymmetrique Activity Index_Medium	2.4765	0.211	11.716	0.000	2.062	2.891
Last Notable Activity_Email Link Clicked	-0.6858	0.338	-2.032	0.042	-1.347	-0.024
Last Notable Activity_SMS Sent	1.2577	0.131	9.606	0.000	1.001	1.514
Occupation_Working Professional	1.4717	0.155	9.484	0.000	1.168	1.776
=====						

As seen above below are the top 3 variables irrespective of numeric/categorical basis their co-efficients,

- Asymmetrique Activity Index\_Low
- Lead Origin\_Lead Add Form
- Last Activity\_Had a Phone Conversation

Below are the top 3 variables basis their co-efficients,

- Asymmetrique Activity Score
- Total Time Spent on Website
- Do not Email (negatively correlated)

- 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 variables that would need focus,

- Asymmetrique Activity Index\_Low
- Lead Origin\_Lead Add Form
- Last Activity\_Had a Phone Conversation

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**

The interns should focus on the below categories to maximize their lead,

- Focus on customers who have Activity Index Low & Medium
  - Lead Origin of cases which has '**Lead Add Form**'
  - Ensure atleast there has some kind of phone conversation for customers for whom lead is there thus making last activity as having phone call with the company resolving their outstanding queries or doubts
  - Customers who have not selected Do Not Email
  - Customers who have spent significantly high on the website
  - Lead source coming from Welingak websites
  - Not focus on customers who have unsubscribed & whose emails have bounced
  - Focus on Working professionals more
  - Not focus People who have been interacting via Olark chats
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**

- Publish creative, educational and informative content for their email subscribers thus increases the curiosity for them to increase activity on the site and raise queries via call backs & chats
- Increase their marketing via Welingak websites
- Ensure more focus and provide as much relevant content for Lead Origin via '**Lead Add Form**'
- Focus on why customers are unsubscribing from the email and try to resolve any issues related to content, experience or other customer service issues
- Campaigns for Working Professionals