

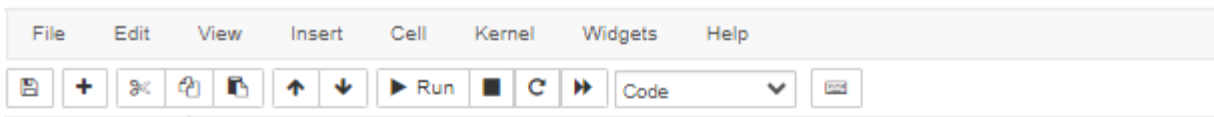
1. Which are the top three variables in your model which contribute most towards the probability of a lead getting converted?

Answer:

Following three variables contribute most towards the probability of getting a lead to be converted and have maximum co-relation coefficient

- 1.) **TotalVisits = 11.1489**
- 2.) **Total Time Spent on Website = 4.4223**
- 3.) **Lead Origin_Lead Add Form = 4.2051**

jupyter DS55_Lead_Scoring_Case_Study_Solution Last Checkpoint: 3 hours ago (autosaved)



Out[81]:

Generalized Linear Model Regression Results

Dep. Variable:	Converted	No. Observations:	4461
Model:	GLM	Df Residuals:	4449
Model Family:	Binomial	Df Model:	11
Link Function:	Logit	Scale:	1.0000
Method:	IRLS	Log-Likelihood:	-2079.1
Date:	Sun, 15 Oct 2023	Deviance:	4158.1
Time:	19:09:00	Pearson chi2:	4.80e+03
No. Iterations:	7	Pseudo R-squ. (CS):	0.3642
Covariance Type:	nonrobust		

	coef	std err	z	P> z	[0.025	0.975]
const	0.2040	0.196	1.043	0.297	-0.179	0.587
TotalVisits	11.1489	2.665	4.184	0.000	5.926	16.371
Total Time Spent on Website	4.4223	0.185	23.899	0.000	4.060	4.785
Lead Origin_Lead Add Form	4.2051	0.258	16.275	0.000	3.699	4.712
Lead Source_Olark Chat	1.4526	0.122	11.934	0.000	1.214	1.691
Lead Source_Welingak Website	2.1526	1.037	2.076	0.038	0.121	4.185
Do Not Email_Yes	-1.5037	0.193	-7.774	0.000	-1.883	-1.125
Last Activity_Had a Phone Conversation	2.7552	0.802	3.438	0.001	1.184	4.326
Last Activity_SMS Sent	1.1856	0.082	14.421	0.000	1.024	1.347
What is your current occupation_Student	-2.3578	0.281	-8.392	0.000	-2.908	-1.807
What is your current occupation_Unemployed	-2.5445	0.186	-13.699	0.000	-2.908	-2.180
Last Notable Activity_Unreachable	2.7846	0.807	3.449	0.001	1.202	4.367

- 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:

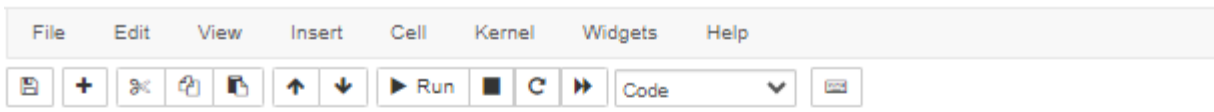
Following three categorical/dummy variables contribute most towards the probability of getting a lead to be converted and have maximum co-relation coefficient

1.) Lead Origin_Lead Add Form = 4.2051

2.) Lead Source_Olark Chat= 1.4526

3.) Lead Source_Welingak Website = 2.1526

jupyter DS55_Lead_Scoring_Case_Study_Solution Last Checkpoint: 3 hours ago (autosaved)



Out[81]:

Generalized Linear Model Regression Results

Dep. Variable:	Converted	No. Observations:	4461
Model:	GLM	Df Residuals:	4449
Model Family:	Binomial	Df Model:	11
Link Function:	Logit	Scale:	1.0000
Method:	IRLS	Log-Likelihood:	-2079.1
Date:	Sun, 15 Oct 2023	Deviance:	4158.1
Time:	19:09:00	Pearson chi2:	4.80e+03
No. Iterations:	7	Pseudo R-squ. (CS):	0.3642
Covariance Type:	nonrobust		

	coef	std err	z	P> z	[0.025	0.975]
const	0.2040	0.196	1.043	0.297	-0.179	0.587
TotalVisits	11.1489	2.665	4.184	0.000	5.926	16.371
Total Time Spent on Website	4.4223	0.185	23.899	0.000	4.060	4.785
Lead Origin_Lead Add Form	4.2051	0.258	16.275	0.000	3.699	4.712
Lead Source_Olark Chat	1.4526	0.122	11.934	0.000	1.214	1.691
Lead Source_Welingak Website	2.1526	1.037	2.076	0.038	0.121	4.185
Do Not Email_Yes	-1.5037	0.193	-7.774	0.000	-1.883	-1.125
Last Activity_Had a Phone Conversation	2.7552	0.802	3.438	0.001	1.184	4.326
Last Activity_SMS Sent	1.1856	0.082	14.421	0.000	1.024	1.347
What is your current occupation_Student	-2.3578	0.281	-8.392	0.000	-2.908	-1.807
What is your current occupation_Unemployed	-2.5445	0.186	-13.699	0.000	-2.908	-2.180
Last Notable Activity_Unreachable	2.7846	0.807	3.449	0.001	1.202	4.367

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:

- 1.) Target almost all the leads that visit the website of X Education during this period
 - 2.) Target leads those spend more time exploring the website of X Education
 - 3.) Target the leads who are visiting the website repeatedly based on no. pages visited per view
 - 4.) Leads generated through referencing.
 - 5.) Target the students as students with industry experience are now being preferred in the industry.
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:

- 1.) Do not focus on students, since they are already studying and would not be willing to enroll into a course specially designed for working professionals, so early in the tenure
- 2.) Do not focus on unemployed leads. They might not have a budget to spend on the course