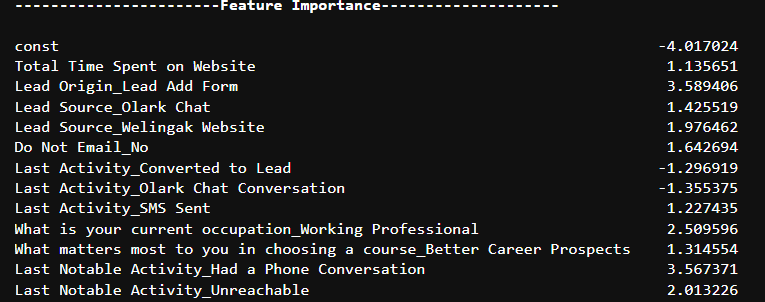
1. Which are the top three variables in your model which contribute most towards the probability of a lead getting converted?

**Ans:** The top three variables in our model which contribute the most towards the lead getting converted are:

* Lead Origin
* Last Notable activity
* What is your current Occupation

The following table shows the importance of each variable wrt their coefficients:

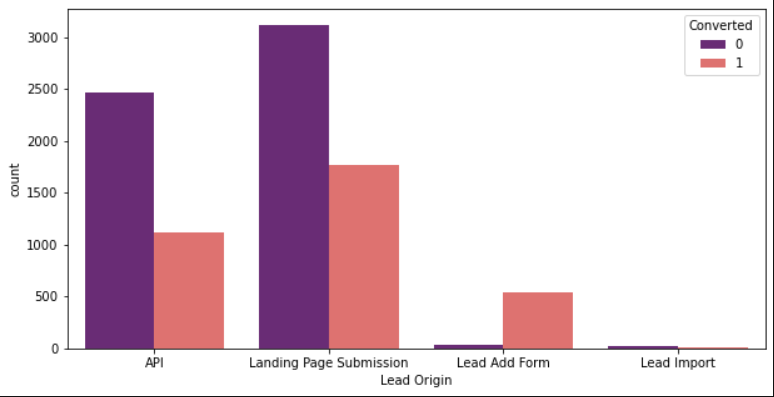


1. What are the top 3 categorical/dummy variables in the model which should be focused the most on in order to increase the probability of lead conversion?

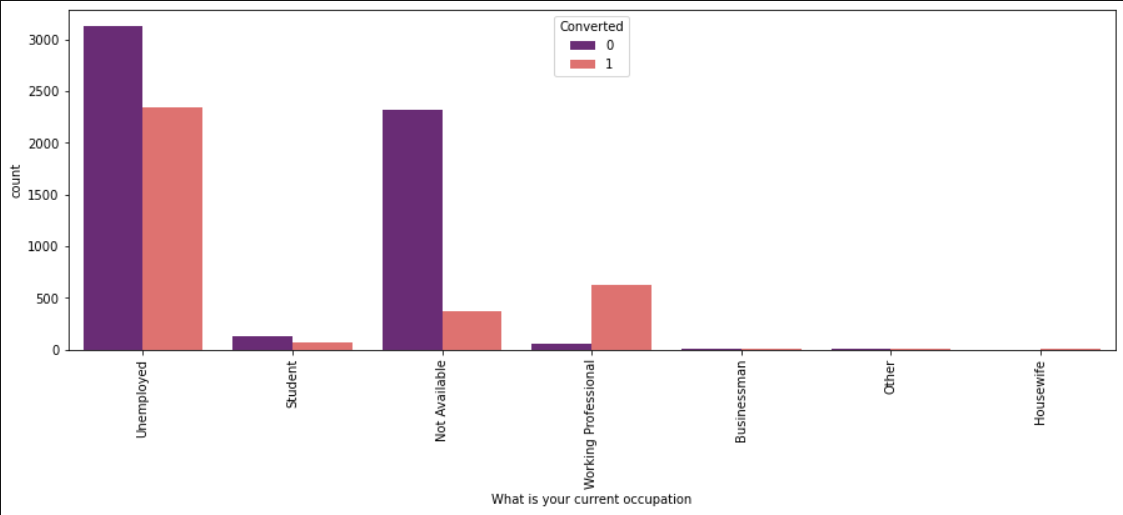
**Ans:** Upon reviewing the final model and which is evident from the EDA the top 3 categorical/dummy variables in which the company should focus to increase the lead conversion are:

* Lead Origin – Lead Add Form
* Last Notable Activity – Had a phone conversation
* What is your current occupation – Working Professional.

**Lead Origin**



**What is your current Occupation**

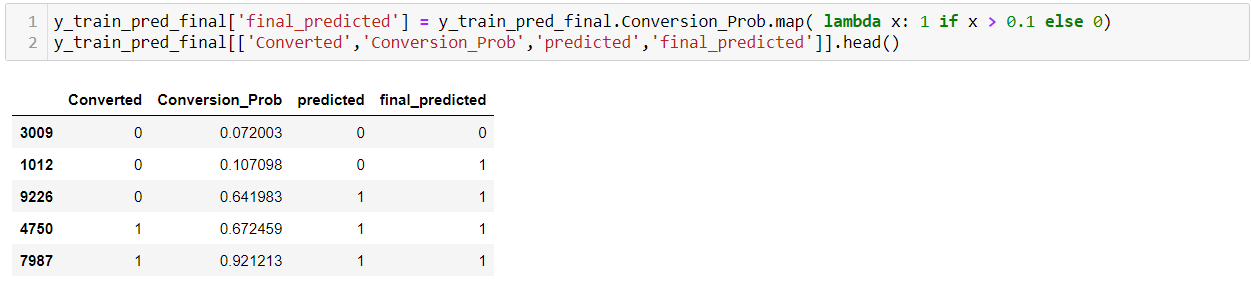


1. X Education has a period of 2 months every year during which they hire some interns. The sales team, in particular, has around 10 interns allotted to them. So, during this phase, they wish to make the lead conversion more aggressive. So, they want almost all of the potential leads (i.e. the customers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

**Ans:** Sensitivity can be defined as the ratio of total number of actual Conversions correctly predicted to the total number of actual conversions. Similarly, Specificity can be defined as the ratio of total number of actual non-conversions correctly predicted to the total number of actual non-conversions. For a model, as one increases, the other decreases and vice versa.

Since the X education has more number of sales team for 2 months and they wish to make the lead conversion more aggressive by wanting almost all the potential leads, we can choose a lower threshold value for Conversion Probability. This will ensure the Sensitivity rating is very high which in turn will make sure almost all leads who are likely to Convert are identified correctly and the agents can make phone calls to as much of such people as possible.

Here we have taken a threshold of 0.1 such that all the conversions i.e ‘1’ are correctly predicted as converted and which also includes some of the non-conversions as converted as this might reduce the conversion rate but increases the conversion count and eventually increasing the revenue of the company.



Accuracy of the train model : 0.604

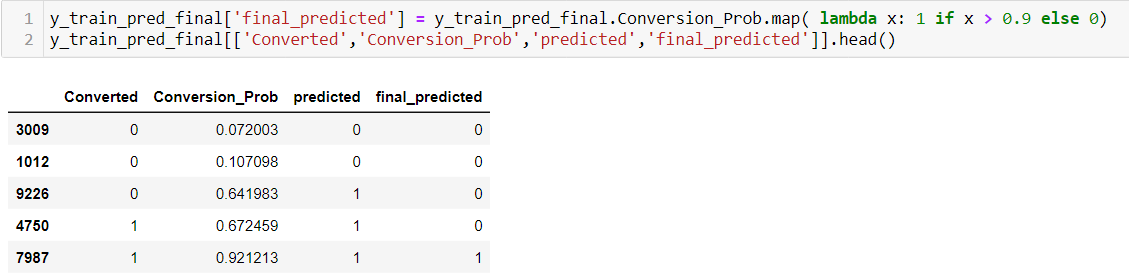
Sensitivity of the train model: 0.986

Specificity of the train model: 0.365

1. 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.

**Ans:** Since X Education has already reached its target for a quarter and doesn’t want to make phone calls unless it is extremely necessary, i.e. they want to minimize the rate of useless phone calls. We can choose a higher threshold value for Conversion Probability. This will ensure the Specificity rating is very high, which in turn will make sure almost all leads who are on the brink of the probability of getting Converted or not are not selected. As a result, they won’t have to make unnecessary phone calls and can focus on some new work. In this way the efficiency of sales team would increase as the conversion rate would be high.

In order to achieve this we have taken a threshold of 0.9 by which we can reduce the number of phone calls and calling to only those leads who have a very high lead score. We can look at the sensitivity and specificity scores below which shows high specificity and low sensitivity.



Accuracy of the train model : 0.713

Sensitivity of the train model: 0.276

Specificity of the train mode : 0.986