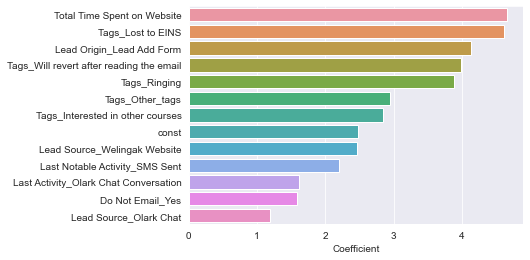
1. Which are the top three variables in your model which contribute most towards the probability of a lead getting converted?

**Answer:**

The top 3 variables in model are:

* Total Time Spent on Website
* Tags Lost to EINS
* Lead Origin Lead Add Form



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?

**Answer:**

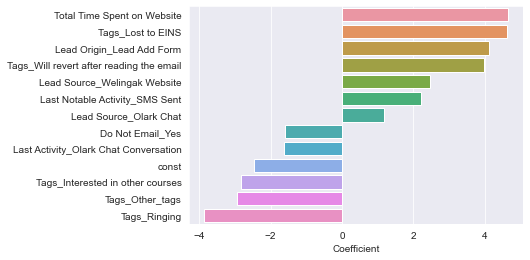
Top 3 Categorical/dummy variables in the model are:

* Tags Lost to EINS
* Lead Origin Lead Add Form
* Tags Will revert after reading the email

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.

**Answer:**

* As they are hiring some interns, the number of persons allotted for the sales team has increased by 10 members. So in our model, we should decrease the cutoff value slightly so that the number of leads will get increased.
* For the above strategy, the most focused variables are mentioned in the figure below along with accuracy, precision and recall.



If the cutoff is decreased to 0.20, then the number of leads for the given data is:

* For train data Hot leads increased to 3355 so that the interns can make more calls to increase the leads conversion
* For test data hot leads increased to 826.
* For train data
  + Precision is 78%
  + Recall is 95 %
  + Accuracy is 87.7 %,
* For test data :
  + Precision is 76%
  + Recall is 94%
  + Accuracy is 87 %

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.

**Answer:**

* As the sales team focuses on some new work and unable to make more calls, the cutoff in our model should be increased based on business constraints so that the number of hot leads will be decreased. (approximately lead score > 80)
* For train data there are 2135 hot leads out of 7238 rows of data where the probability is more than 0.8.
* For test data there are 534 leads out of 1810 rows where the probability is more than 0.8.
* We can call these leads whenever it is extremely necessary to explain Education courses to increase the conversion rate.
* For Train data
  + Precision: 76 %
  + Recall: 75%
  + Accuracy is 89.2 %
* For test data
  + Precision : 95 %
  + Recall : 76 %
  + Accuracy : 90 %