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

|  |  |
| --- | --- |
| Features | coeff |
| Tags\_Will revert after reading the email | 1.512422 |
| Lead Origin\_Lead Add Form | 1.298576 |
| Tags\_Ringing | 1.187626 |

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?

|  |
| --- |
| 'Lead Origin', |
| 'Tags', |
| 'Last Notable Activity' |

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.

In the above scenario, X Education needs maximum correct Lead Predictions. Hence, we should focus on maximum Precision.

Reason:

Precision is the metric which tells you the proportion of positive identifications was actually correct?

A model that produces no false positives has a precision of 1.0.

Precision = TP / (TP+FP)

Our final model has Precision Score of 89.55% ≈ 90%

That means when it predicts a Lead is Converted, it is correct 90% of the time.

In order to ensure this the **interns can** **Make as many phone calls as possible to get maximum Conversions into /Hot Leads** to following groups of Customers

… based on time spent on Website if Maximum

… Multiple visits to websites by same customer

… Their last activity is through SMS or Olark Chat conversation

… if they are unemployed or Working Categories.

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.

In the above scenario, X Education needs no Wrong Lead Predictions. Hence, we should choose a model with maximum Recall.

Reason:

Recall is the metric which tells you the proportion of actual positives was identified correctly

A model that produces no false negatives has a Recall of 1.0.

Precision = TP / (TP+FN)

Our final model has Precision Score of 82.6 % ≈ 83%

in other words, it correctly identifies 83% of all Leads.

Our model tells ‘Tags Will revert after reading the email’ is a major feature deciding the Lead Conversion … hence the **Intern team should target maximum emails to be sent to potential leads to find Hot Leads**

This strategy can be used to only those customers that have very high chances of buying the course