Summary

This analysis is done for X Education and to find ways to get more industry professionals to join their course. The basic data is to provide us how potential customers visit the site how much time they spend, and how they reach to the site and the conversion rate.

The following are the steps which are used In this assignment:

1. Cleaning Data

The data was partially clean except for a few null values and option select had to be replaced with a null value since it did not give much information. Few of the values were changed to 'not provided' so as to not lose data. Although they were later removed while making dummies.

2. EDA

A EDA was done to check the condition of our data. It was found that a lots of elements in our data where removed.

3. Dummy Variables

The dummy variables were created and elements where removed.

4. Train-Test Split

The split was done at 70% and 30% for train data and test data

5. Model Building

Firstly, RFE was done to attain the top 15 relevant variables. Later the rest of the variables were removed manually depending on the VIF values and p-value(the variables with VIF <5 and p-value <0.05 were kept)

6. Model Evaluation

A confusion matrix was made, later on the optimum cutoff value (using ROC curve) was used to find the accuracy, sensitivity and specify which came to be around 80% each.

7. Prediction

Prediction was done on the test data frame and with an optimum cut off as 0.35 with accuracy, sensitivity and specificity of 80%

8. Precision- Recall

This method is used to recheck the cutoff of 0.41 which was found with precision around 73% and recall around 75% on the test data frame

It was found that the variable that mattered the most in the potential buyers are(In descending order)

1. Total time spent on website
2. Total number of visits
3. When the lead source was
   a. Google
   b. Direct Traffic
   c. Organic Search
   d. Welingak website
4. When the last activity was
   a. SMS
   b. Olark chat conversation
5. When the lead origin is lead add format
6. When their current occupation is a working professional

Keeping in mind X education can flourish as they have a very high chance to get almost all the potential buyers to change their mind and buy their course.