Summary

This analysis is done for X Education and to find ways to get more industry professionals tojoin their courses. The basic data provided gave us a lot of information about how thepotential customers visit the site, the time they spend there, how they reached the site and theconversionrate.

Thefollowing are the steps used:

# Cleaningdata:

The data was partially clean except for a few null values and the option select had tobe replaced with a null value since it did not give us much information. Few of thenull values were changed to ‘not provided’ so as to not lose much data. Although theywere later removed while making dummies. Since there were many from India andfew from outside, the elements were changed to ‘India’, ‘Outside India’ and ‘notprovided’.

# EDA:

A quick EDA was done to check the condition of our data. It was found that a lot ofelements in the categorical variables were irrelevant. The numeric values seems good and no outliers were found.

# DummyVariables:

The dummy variables were created and later on the dummies with ‘not provided’elements were removed. Fornumeric values weused the MinMaxScaler.

# Train-Testsplit:

Thesplit was done at70% and 30%for trainandtest datarespectively.

# ModelBuilding:

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

# ModelEvaluation:

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

It was found that the variables that mattered the most in the potential buyers are (Indescending order):

1. Thetotaltime spendon theWebsite.
2. Totalnumber of visits.
3. Whentheleadsourcewas:
   1. Google
   2. Directtraffic
   3. Organicsearch
   4. Welingakwebsite
4. Whenthelastactivitywas:
   1. SMS
   2. Olarkchatconversation
5. Whentheleadorigin isLeadaddformat.
6. Whentheircurrent occupationis asa working professional.

Keeping these in mind the X Education can flourish as they have a very high chance to getalmostall thepotential buyers tochange their mindand buy theircourses.

X X X X