

SUMMARY OF CASE STUDY

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

The following are the steps used:

1. Importing and Cleaning Data: Firstly import the data and then the data was clean om some parameters also imputing the missing values.
2. EDA: A EDA was done to check the condition of our data. It was found that a lot of elements in the categorical variables were irrelevant.
3. Dummy Variables: The dummy variables were created and later on the dummies with 'not provided' elements were removed.
4. Train-Test split
5. Model Building: Firstly, RFE was done to attain the top 18 relevant variables. Later the rest of the variables were removed manually depending on the VIF values and p-value
6. Model Evaluation: 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 81% each.
7. Prediction: Prediction was done on the test data frame and with accuracy, sensitivity and specificity of 81%.
8. Precision: This method was used to recheck and a cut off.

The Most Potential Buyers are:

1. The total time spend on the Website.
2. Total number of visits.
3. When the lead source was from Google ,Direct traffic, website and sms.
4. When their current occupation is as a working professional.

Keeping these in mind the X Education can have a very high chance to get almost all the potential buyers to change their mind and buy their courses.