

Summary

This analysis is done for X Education and to find ways to get more industry professionals to join their courses. 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) Cleaning Data:
 - i) While submitting detail by the customer/visitors, he may not select any value in particular field. In that condition it contains 'Select'. Which is not the right value. So we changed it to Null value.
 - ii) Dropped columns which contains only one unique value. As it won't affect our analysis.
 - iii) Dropped columns having high number of null values
 - iv) Categorization of column 'Country'.
 - v) Drop rows having null values in any column.
- 2) EDA(Exploratory data analysis):

Performed EDA for univariate analysis in Categorical variables and numerical variables
Check if any outliers are there.
- 3) Dummy variable creation:
 - i) Total 9074 rows and 22 columns to be Analyzed
 - ii) Dummy Variables are created for Object type variables
- 4) Model Building:
 - i. Splitting the data into testing and training set. Split the data into 75:25 ratio.
 - ii. Use RFE for feature selection
 - iii. Selected top 15 variables from RFE
 - iv. Building Model by removing the variable whose p- value is greater than 0.05 and vif value is greater than 5
 - v. Predictions on test data set
 - vi. Accuracy is around 79%
- 5) ROC Curve:
 - i. The area under ROC curve is 0.87 which is a very good value.
 - ii. optimal cut off is at 0.35.
- 6) Conclusion:

It was found that the variables that mattered the most in the potential buyers are :

 - i. The total time spend on the Website.
 - ii. Total number of visits.
 - iii. When the lead source was:
 - a. direct traffic
 - b. welingak website
 - iv. When the last activity was Olark chat conversation.
 - v. When the lead origin is Lead add format.

- vi. When their current occupation is as a working professional.
Keeping these in mind the 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 courses.