

SUMMARY

This analysis is done for an Education company X Education 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, how they have reached the site and the conversion rate.

The following are the steps used:

1. Data Cleaning:

The data was partially clean except for a few null values and the option select had to be replaced with a null value since it did not give us much information.

2. EDA:

A quick EDA was done to check the condition of our data. It was found that a lot of elements in the categorical variables were irrelevant. The numeric values seems good and Outliers were found in Total Visits.

3. Dummy Variables:

The dummy variables were created and later on the dummies with 'not provided' elements were removed.

4. Train-Test split:

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

5. Model Building:

- Firstly, RFE was done to attain the top 20 relevant variables.
- Later the rest of the variables were removed manually depending on the VIF values and p-value .
- All the VIF values were good and all the p-values were below 0.05.

6. Model Evaluation:

A confusion matrix was made. Later on the optimum cut off value i.e. 0.31 using ROC curve was used to find the accuracy, sensitivity which came to be around 81% and specificity which came to be around 81% each.

7. Prediction:

Prediction was done on the test data frame and with an optimum cut off as 0.41 with accuracy of 81%, sensitivity around 82% and specificity of around 80%.

8. Precision – Recall:

This method was also used to recheck and a cut off of 0.41 was found with Precision around 76% and recall around 75% on the test data frame.

9. Conclusion:

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

1. The TotalVisits
2. The Total time spend on the Website.
3. Lead Origin_Lead Add Form
4. Lead Source :
 - Direct Traffic
 - Google
 - Organic Search
 - Referral Sites
 - Welingak Website
5. Do not Email : Yes
6. When the Last Activity was:
 - Email Bounced
 - Olark Chat Conversation

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.