

Summary

This case study is done for X Education. The aim is to find ways to get more people to join their courses. The basic data provided contains information about the people visiting their site, time spent on website, conversion rate etc.

We carried out the following steps for our analysis:

1. Cleaning data:

Certain columns were removed which were not useful for analysis. Few of the null values were changed accordingly so as to not lose much data. Although they were later removed while making dummies. Since there were many from India and few from outside, we dropped the column country.

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, so we dropped those columns. No outliers were found.

3. Dummy Variables:

The dummy variables were created and later some columns whose dummy was already created were dropped.

4. Train-Test split:

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

5. Model Building:

We used RFE for feature selection with top 17 variables as the required output. After removing the features with p value $> .05$ we checked the VIF and it was < 5 for all the features.

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, specificity, precision and recall. They were 77%, 83%, 73%, 66% and 83% respectively.

7. Prediction:

Prediction was done on the test data frame and with an optimum cut off as 0.3 with accuracy, sensitivity and specificity of 80%.

8. Precision – Recall:

This method was also used to recheck and a cut off of 0.3 was found with Precision around 65% and recall around 66% on the test data frame.

- The customer who fills the form are the potential leads.
- We must majorly focus on working professionals.
- If the lead source is referral, they may not be the potential lead.
- We must majorly focus on leads whose last activity is Email opened or SMS sent.
- It's better to focus least on customers to whom the sent mail is bounced back.
- It's always good to focus on customers, who have spent significant time on our website.

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.