

Summary

We conducted an analysis on X Education to identify strategies for attracting more industry professionals to their courses. Using the provided data, we were able to gain insight into customer behaviour related to course selection, website engagement, and conversion rates. We also learned about the demographics of our target audience, including whether they are students or working professional.

The following steps were used:

Data Cleaning :

The data was inspected and cleaned partially and the not selected options were replaced with null values as it did not provide much information. The null values were imputed with 'not provided' so that not much of data is lost. The same was later removed while assigning dummy.

EDA :

EDA steps were followed on the data and univariate, multivariate analysis were carried out. Since majority of categorical variables were irrelevant they were not included in the analysis. No outliers were found in the process.

Dummy Variable :

The dummy variable was created and MinMaxScaler was used for numeric values.

Model Building:

To identify the most relevant variables, we used a feature selection method called recursive feature elimination (RFE). This helped us narrow down the list to the top 15 variables. We then manually removed additional variables based on their variance inflation factor (VIF) and p-value. Specifically, we kept only those variables with a VIF less than 5 and a p-value less than 0.05.

Model Evaluation :

Confusion matrix was used. Later optimum cut off value was used to find the accuracy, sensitivity and specificity which was around 80%.

Precision- Recall :

This method was used to recheck and a cutoff of 0.35 was observed with Precision 73% and Recall around 76% on test data.

It was observed that the variables with high impact are :

- Total time spend on Website
- Total number of visits
- Lead source was Google, Welingak Website
- When current Occupation is a working professional.
- With the above factors X education can covert potential leads and could gain profit.