

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.

Summary of Steps Taken for Analysis by X Education:

Data Cleaning:

Addressed null values and irrelevant options, replacing them with appropriate labels.

Transformed elements related to location to 'India', 'Outside India', and 'not provided' categories.

Exploratory Data Analysis (EDA):

Conducted a quick exploration to assess data quality.

Identified irrelevant elements in categorical variables but found numeric values satisfactory without outliers.

Dummy Variables:

Created dummy variables and removed those with 'not provided' elements.

Applied MinMaxScaler to numeric values for normalization.

Train-Test Split:

Split the data into 70% for training and 30% for testing.

Model Building:

Employed Recursive Feature Elimination (RFE) to select top 15 relevant variables.

Manually removed additional variables based on VIF values and p-values (kept variables with $VIF < 5$ and $p\text{-value} < 0.05$).

Model Evaluation:

- Constructed a confusion matrix.
- Determined the optimal cutoff value using the ROC curve to assess accuracy, sensitivity, and specificity, achieving approximately 80% for each.

Prediction:

Made predictions on the test data using an optimal cutoff value of 0.35, yielding 80% accuracy, sensitivity, and specificity.

Precision-Recall:

Utilized precision-recall to verify results, identifying a cutoff of 0.41 with precision around 73% and recall around 75% on the test data.

Key Findings:

Identified top variables influencing potential buyers, including total time spent on the website, total number of visits, lead source (Google, Direct traffic, Organic search, Welingak website), last activity (SMS, Olark chat conversation), lead origin (Lead add format), and current occupation (working professional).

These insights can guide X Education in focusing efforts on potential buyers, thereby increasing course enrollment rates.

By leveraging these findings, X Education can effectively target and engage potential buyers, thereby improving their conversion rates and overall business success.
