

This analysis is done for X Education to attract more industry professionals to enroll in their courses. The provided data revealed insights into customer site visits, duration spent on the site, referral sources, and conversion rates. The steps involved were:

- **Data Cleaning:** The data was mostly clean, with a few null values addressed. The option 'select' was replaced with null as it provided little information. Some null values were labeled 'not provided' to retain data. Later, 'India', 'Outside India', and 'not provided' categories were standardized.
- **Exploratory Data Analysis (EDA):** A quick EDA highlighted irrelevant elements in categorical variables. Numeric values appeared sound with no outliers detected.
- **Dummy Variables:** Dummy variables were created, and those with 'not provided' were removed. Numeric values were scaled using MinMaxScaler.
- **Train-Test Split:** Data was split into 70% for training and 30% for testing.
- **Model Building:** RFE identified the top 15 variables, and others were removed based on VIF values (<5) and p-values (<0.05).
- **Model Evaluation:** A confusion matrix was generated. Optimal cutoff values from the ROC curve resulted in 80% accuracy, sensitivity, and specificity.
- **Prediction:** Test data was predicted using an optimal cutoff of 0.35, achieving 80% accuracy, sensitivity, and specificity.
- **Precision-Recall:** Precision-recall analysis indicated a cutoff of 0.41 with 73% precision and 75% recall on test data.
- **Key Factors for Potential Buyers (Descending Order):**
 1. Total time spent on the Website.
 2. Total number of visits.
 3. Lead source:
 4. Google
 5. Direct traffic
 6. Organic search
 7. Welingak website
 8. Last activity:
 9. SMS
 10. Olark chat conversation
 11. Lead origin: Lead add format.
 12. Current occupation: Working professional.

By focusing on these factors, X Education can effectively attract potential buyers and enhance course enrollment rates.