

## Summary of Analysis for X Education

**Objective:** To identify strategies for attracting more industry professionals to enrol in X Education's courses by analysing visitor behaviour and conversion data.

### Methodology:

1. **Data Cleaning:**
  - Addressed null values and irrelevant entries. reclassified geographical data into 'India', 'foreigner', and 'unknown'.
  - Null values were categorized as 'unknown' before being removed for dummy variable creation.
2. **Exploratory Data Analysis (EDA):**
  - Conducted initial EDA to assess data quality. Identified irrelevant categories in categorical variables. Confirmed no significant outliers in numeric data.
3. **Dummy Variables Creation:**
  - Created dummy variables for categorical data. Removed dummies with 'unknown' entries.
4. **Train-Test Split:**
  - Split the data into 72% for training and 28% for testing.
5. **Model Building:**
  - Used Recursive Feature Elimination (RFE) to select the top 15 relevant variables.
  - Manually removed variables with high Variance Inflation Factor (VIF > 5) and non-significant p-values ( $p > 0.05$ ).
6. **Model Evaluation:**
  - Constructed a confusion matrix. Determined the optimum cut-off value using the ROC curve. Achieved an accuracy, sensitivity, and specificity of approximately 80%.
7. **Prediction:**
  - Applied the model to the test data with an optimum cut-off of 0.35, maintaining an accuracy of 77% , sensitivity of 86%, and specificity of 72%.
8. **Precision-Recall Analysis:**
  - Verified results using precision-recall analysis. Identified an optimal cut-off of 0.41, with precision at 76% and recall at 78%.

### Key Findings:

The most influential factors for potential buyers, in descending order of impact, are:

1] *TotalVisits*

2] *Total Time Spent on Website*

3] *Page Views Per Visit*

*4] Lead Origin\_Lead Add Form*

*5] Lead Source\_Olark Chat*

*6] Lead Source\_Welingak Website*

*7] Last Activity\_Email Bounced*

*8] WLast Activity\_SMS Sent*

*9] What is your current occupation\_Working Profes.*

*10] Lead Source\_Olark Chat*

*11] Last Activity\_Email Bounced*

*12] Last Notable Activity\_Unreachable*

*13] Last Notable Activity\_Had a Phone Conversation*

### **Conclusion:**

By focusing on these key factors, X Education can significantly improve its conversion rates and attract a higher number of industry professionals to enroll in their courses. Implementing targeted strategies based on these insights will likely enhance the institution's growth and customer acquisition success.