

Summary

This analysis focuses on X Education's efforts to attract industry professionals to enroll in their courses. The provided dataset offers insights into customer site visits, duration spent on the site, referral sources, and conversion rates. Here's a breakdown of the steps taken:

1. Data Cleaning:

- Null values were handled, with 'Select' options replaced by null to preserve data integrity.
- Some null values were filled with 'Not Specified' to retain data.

2. Exploratory Data Analysis (EDA):

- Initial EDA revealed irrelevant elements in categorical variables.
- Outliers were identified in numeric variables and addressed using the quantile method.

3. Dummy Variables:

- Dummy variables were created, and original elements were subsequently removed.
- Numeric values were standardized using StandardScaler.

4. Train-Test Split:

- The dataset was divided into 70% training and 30% testing sets.

5. Model Building:

- Recursive Feature Elimination (RFE) was employed to select the top 15 relevant variables.
- Variables were further pruned based on VIF values (<5) and p-values (<0.05).

6. Model Evaluation:

- A confusion matrix was generated, and the optimal cut-off value was determined using the ROC curve.
- Accuracy, sensitivity, and specificity were calculated, with the model achieving 82%.

7. Prediction:

- Predictions were made on the test dataset, with an optimal cut-off of 0.374, resulting in improved accuracy, sensitivity, and specificity.

8. Precision-Recall Analysis:

- Precision and recall scores were calculated, yielding values of 73 and 80, respectively.

Key Findings:

- Significant variables influencing potential buyers include total time spent on the website, total number of visits, and specific lead sources like Google, Direct traffic, Organic search, and Welingak website.
- Last activity, particularly SMS and Olark chat conversation, also plays a crucial role.
- Lead origin and occupation as a working professional emerged as key factors in attracting potential buyers.

Conclusion:

By focusing efforts on the identified influential variables, X Education stands to significantly enhance its conversion rates and attract a broader pool of potential buyers. The precision and recall scores validate the effectiveness of the model in targeting high-conversion leads, paving the way for improved business outcomes.