Lead Scoring Summary

Analysis Summary for X Education:

Introduction

The analysis focuses on enhancing the enrollment of industry professionals in X Education courses. The dataset provided insights into customer site visits, duration, referral sources, and conversion rates.

Technical Steps

1. **Data Cleaning:**

   - Removed redundant features.

   - Replaced the 'Select' option with null values.

   - Dropped columns with over 40% null values.

   - Identified and dropped highly skewed columns.

   - Imputed missing values using appropriate aggregate functions (mean, median, mode).

   - Detected outliers.

2. **Exploratory Data Analysis (EDA):**

   - Reviewed data quality; observed irrelevant elements in categorical variables.

   - Identified outliers in numeric values.

   - Conducted Univariate and Bivariate Analysis to understand variable distributions and their relationships with the target.

3. **Dummy Variables:**

   - Created dummy variables for categorical columns.

4. **Scaling:**

   - Applied Standard Scalar to scale continuous variables.

5. **Train-Test Split:**

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

6. **Model Building:**

   - Utilized Recursive Feature Elimination (RFE) for the top 20 relevant variables.

7. **Model Evaluation:**

   - Constructed a confusion matrix.

   - Determined an optimal cut-off value using the ROC curve for accuracy, sensitivity, and specificity (80%).

8. **Prediction:**

   - Performed predictions on the test dataset with an optimum cut-off at 0.37, achieving 80% accuracy, sensitivity, and specificity.

9. **Precision-Recall:**

   - Applied Precision-Recall to recheck, resulting in a cut-off value of 0.41.

10. **Conclusion:**

    - Identified critical variables influencing potential buyers:

      - Total time spent on the website.

      - Total number of visits.

      - Lead sources like Olark Chat.

      - Last activities such as SMS and Olark chat conversations.

Conclusion

The analysis pinpointed key factors impacting potential buyer engagement, providing insights crucial for strategizing efforts to attract industry professionals to X Education's courses.