

Lead Score Case Study

Summary Report

Analysis of the Lead Scoring Case Study

1. Data Inspection

- **Initial Data Overview:** The dataset was examined for structure, missing values, and basic statistics.

2. Data Cleaning

- Handling Missing Values and Outliers Detection and Treatment.

3. Data Analysis

- Univariate Analysis and Bivariate Analysis

4. Data Preprocessing

1. **Splitting Data:** The dataset is split into training and testing sets using a 70:30 ratio.
2. **Feature Scaling:** Applied StandardScaler to standardize features such as 'Total Visits,' 'Total Time Spent on Website,' and 'Page Views per Visit.'

5. Correlation Matrix

- A heatmap is used to identify multicollinear variables and variables highly correlated with the target variable ('Converted').

6. Model Building

1. **Logistic Regression:** Implemented using sklearn's Logistic Regression.
2. **Feature Selection:** Recursive Feature Elimination (RFE) is used to select the top features.
3. **Model Building:** The final logistic regression model is defined with significant variables contributing to lead conversion.

7. Model Evaluation

Confusion Matrix: Used to calculate Accuracy, Sensitivity, Specificity, Precision, and Recall for both training and testing datasets.

Recommendations Based on the Analysis

1. **Maximize Website Engagement:**
 - Encourage leads to spend more time on the website through engaging content, interactive tools, and valuable resources.
2. **Optimize Lead Origin Channels:**
 - Focus marketing efforts on lead origin sources that have shown higher conversion rates.

3. **Improve Email Communication:**

- Develop strategies to reduce the 'Do Not Email' rate by ensuring email content is relevant, personalized, and valuable to leads.

4. **Refine Lead Source Strategies:**

- Concentrate efforts on lead sources that contribute positively to conversion rates.

5. **Monitor and Optimize Website Navigation:**

- While 'Page Views Per Visit' has a negative impact, it could be due to a complex or confusing website layout. Simplify navigation to improve user experience.

6. **Analyze and Enhance Last Activity:**

- Investigate what activities are beneficial and promote those to increase conversion rates.

7. Make it mandatory to fill the below columns so that more information can be extracted based on these:

- Specialization
- How did you hear about X Education
- Lead Profile
- City

