Lead Scoring Case Study

Summary:

This study is conducted for X Education with the aim of attracting more industry professionals to enroll in their courses. The fundamental data provided has yielded valuable insights into the browsing behaviour of potential customers, including their site visit patterns, duration of engagement, referral sources, and the overall conversion rate.

The following are the steps used:

1. Read and comprehend the data: Load the dataset, read its contents, and analyze its structure or shape.
2. Data Cleaning: Identify and handle null values and specific values. Replace the designated value with "nan" to signify missing information. Remove columns with a high percentage of null values and impute missing values using median/mode for numerical variables. Create a new categorical variable for categorical features as required.
3. Exploratory Data Analysis (EDA): Conducted an in-depth analysis, revealing that numerous elements within categorical variables were irrelevant. The numeric values appeared satisfactory, exhibiting no outliers.
4. Dummy Variables: Created dummy variables to represent categorical values in the dataset.
5. Train-Test Split: The dataset was divided into training and testing sets, with a distribution ratio of 70% for training and 30% for testing.
6. Feature Scaling: Applied the MINMAXScaler to normalize the original numerical values in the dataset.

7. Model Building: Implemented Recursive Feature Elimination to identify the top 15 most relevant variables. Subsequently, based on VIF values and p-values, removed the remaining variables
8. Model Evaluation: A confusion matrix was generated, and subsequently, the optimal cutoff value (determined using the ROC curve) was employed to calculate accuracy, sensitivity, and specificity, resulting in an approximately 80% accuracy.
9. Prediction: Utilized an optimal cutoff of 0.35 for predictions on the test data frame, achieving an accuracy, sensitivity, and specificity of 80%.

Discussion:

To enhance lead conversion during the intern-hiring period, X Education can adopt a targeted strategy by focusing on high-potential leads, specifically those from the Welingak Website and reference sources. The sales team should prioritize calls to these leads and leverage effective communication channels, such as SMS and email, to improve conversion rates. Additionally, prioritizing leads based on website engagement can further increase the likelihood of success.

To minimize unnecessary phone calls post-achieving quarterly sales targets ahead of schedule, the sales team can shift focus to lead nurturing activities like personalized emails, SMS messages, and targeted newsletters. Automation of SMS messages for high-conversion probability customers can optimize outreach efforts. Seeking feedback from existing customers helps refine lead quality and conversion rates. Collaboration among sales, management, and data scientists is crucial to continually improve the lead conversion model. Developing a strategy for discounts or incentives can motivate potential customers to take prompt

action. This comprehensive approach ensures efficient lead conversion and minimizes unnecessary outreach efforts after reaching sales targets.