Summary Report

Lead Score Case Study

Context & Objective:

X Education, a provider of online courses for industry professionals, faces challenges with its lead conversion rate, which currently hovers around 30%. With a daily influx of leads from various sources such as website visits and referrals, the sales team requires a strategic approach to identify high-potential leads, or "Hot Leads." The objective of this study was to develop a logistic regression model to assign lead scores, enabling a data-driven prioritisation of leads and aiming to increase the conversion rate to around 80%.

Key Methodology:

- **Data Preparation:**
  - The dataset was reviewed and cleaned, with null values appropriately handled to preserve data integrity.
  - Categorical variables were encoded, and numeric features were scaled using MinMaxScaler to ensure compatibility with the model.

- **Exploratory Data Analysis (EDA):**
  - A comprehensive analysis was conducted to understand customer behavior and variable relationships.
  - Irrelevant features were flagged for exclusion, while critical predictors were identified.

- **Feature Engineering:**
  - Dummy variables were created for categorical features to enable their inclusion in the regression model.
  - Non-informative categories and redundant entries were removed for better model performance.

- **Model Building and Optimisation:**
  - Recursive Feature Elimination (RFE) was applied to select the top predictors of lead conversion.
  - Further refinement of features was performed based on Variance Inflation Factor (VIF) and p-values, ensuring statistical and practical significance.

- **Performance Evaluation:**
  - The model was evaluated using metrics such as accuracy, sensitivity, and specificity.
  - An optimal cutoff value of 0.35 was determined using the ROC curve, yielding balanced performance with approximately 80% for all three metrics.

- **Prediction Validation:**
  - Predictions on the test dataset validated the model's performance, with an optimal cutoff of 0.41 achieving a precision of 73% and recall of 75%.

Insights and Business Impact:

The study highlighted several critical factors driving lead conversion, including:

- Website engagement metrics like total time spent and number of visits.
- Lead sources such as Google, direct traffic, and organic search.
- Customer interaction channels, including SMS and Olark chat.
- Specific attributes such as lead origin (e.g., Lead Ad Form) and occupation as a working professional.

By focusing on these variables, X Education can adopt a targeted approach to prioritise and nurture high-potential leads, ensuring efficient resource allocation and significantly boosting conversion rates.

Recommendations:

- **Streamline Lead Management:** Prioritise leads based on the lead scores generated by the model to optimize sales team efforts.
- **Enhance Website Engagement:** Invest in strategies to increase user engagement on the website, as time spent and visits are key predictors of conversion.
- **Focus on High-Impact Channels:** Strengthen marketing efforts on high-performing lead sources like Google and direct traffic while optimising engagement strategies for SMS and chat interactions.

By implementing these recommendations, X Education is well-positioned to meet its goal of achieving an 80% lead conversion rate and driving growth in its customer base.