LEAD SCORE CASE STUDY REPORT

For this case study, the goal is to build a logistic regression model to assign lead scores to potential customers. The model should predict the likelihood of a lead converting into a paying customer based on various attributes provided in the dataset. Here's a summary of the approach and the key findings:

Problem Statement:
X Education wants to identify the most promising leads likely to convert into paying customers. The company requires a model that assigns a lead score between 0 and 100 to each lead, with higher scores indicating a higher likelihood of conversion.

Analysis Approach:
1. Data Cleaning: Handle missing values and 'Select' levels in categorical variables.
2. Exploratory Data Analysis (EDA): Understand the distribution of variables and their impact on conversion.
3. Feature Engineering: Create new features if needed to improve model performance.
4. Model Building: Build a logistic regression model to predict conversion probability.
5. Model Evaluation: Evaluate the model using metrics like accuracy, precision, recall, and F1-score.
6. Adjusting to Future Requirements: Address additional problems provided by the company to make the model adaptable to future changes.

Business Insights:
- The model can help X Education prioritize leads by focusing on those with higher lead scores, thus increasing the conversion rate.
- Insights from the model can be used to tailor marketing and communication strategies for different lead segments, improving overall lead nurturing efforts.

Future Recommendations:
- Continuously monitor and update the model to adapt to changing market dynamics and customer behaviors.
- Consider incorporating additional data sources or advanced machine learning techniques to improve model performance over time.

Learnings:
- Importance of data preprocessing and feature engineering in improving model accuracy.
- Impact of model performance on business outcomes, emphasizing the need for continuous improvement and adaptation.

Conclusion:
The logistic regression model developed for X Education provides a valuable tool for identifying and prioritizing potential leads. By focusing on leads with higher scores, the company can optimize its lead nurturing efforts and improve conversion rates.

Model Performance:

METRIC	VALUE
Accuracy	0.815
Sensitivity	0.928
Specificity	0.751
Precision	0.680

Submitted by:
ISHAN AGRAWAL
SHYAM SUNDAR
AKANKSHA GUPTA