SUMMARY REPORT

1. Introduction

This project aims to develop a lead scoring model to predict lead conversion probabilities, enhancing targeted marketing strategies. A logistic regression model is used to improve conversion rates by identifying high-potential leads.

2. Data Collection and Preprocessing

- Missing values were handled through imputation and removal.
- Categorical variables were encoded using one-hot encoding.
- Numerical features were standardized for consistency.
- Duplicate and irrelevant features were removed to optimize model performance.

3. Exploratory Data Analysis (EDA)

- Data analysis revealed trends, distributions, and correlations.
- Most leads originated from organic search, direct traffic, and referrals.
- Leads spending more time on the website and visiting multiple pages had a higher conversion probability.
- Variables like "Lead Source" and "Last Activity" significantly influenced conversions.

4. Model Development

- Logistic regression was chosen for its simplicity and interpretability.
- Data was split into training (80%) and testing (20%) sets.
- Feature selection was based on statistical tests and correlation analysis.
- The model was trained and optimized through hyperparameter tuning.
- Performance was evaluated using key metrics like accuracy, precision, and recall.

5. Model Evaluation and Results

The logistic regression model achieved strong predictive performance:

- **Accuracy:** ~85%
- **Precision:** 78%
- **Recall:** 82%
- **F1-score:** 80%
- **ROC-AUC Score:** 0.88

6. Business Insights and Recommendations

- Focus marketing efforts on high-scoring leads for better conversion rates.
- Develop engagement strategies for medium-scoring leads to increase conversions.

- Optimize lead generation channels based on conversion performance.
- Update the model regularly with new data to adapt to trends and maintain accuracy.

7. Conclusion

The lead scoring model effectively identifies high-potential leads, allowing the education platform to enhance marketing strategies and improve conversion rates.