Summary Report

Lead Scoring Analysis for X Education

1. Data Splitting and Scaling:

- Split the dataset into 70% training data and 30% testing data.

- Scaled numerical columns to standardize the data.

2. Feature Selection and Model Building:

- RFE (Recursive Feature Elimination): Used RFE to select the top 15 relevant variables.

- Manual Selection: Further refined features by removing those with high Variance Inflation Factor (VIF > 5) and non-significant p-values (p > 0.05).

3. Model Training:

- Trained the final model on the training set using the selected features.

- Predicted target variables on the training data.

4. Model Evaluation:

- Confusion Matrix: Constructed a confusion matrix to evaluate the model's performance.

- ROC Curve Analysis: Determined the optimal cut-off value using the ROC curve, achieving around 90% for accuracy, sensitivity, and specificity.

- AUC: The area under the ROC curve (AUC) was 0.97, indicating a highly predictive model.

5. Test Data Prediction:

- Applied the model to the test dataset.

- Achieved approximately 92% for accuracy, sensitivity, and specificity on test data.

6. Precision-Recall Analysis:

- Precision: Scored 89% on the test data.
- Recall: Scored 91% on the test data.

7. **Final Observations:**

- **Train Data Performance:**
    - Accuracy: 92.29%
    - Sensitivity: 91.70%
    - Specificity: 92.66%
- **Test Data Performance:**
    - Accuracy: 92.78%
    - Sensitivity: 91.98%
    - Specificity: 93.26%

Implications for X Education:

- **Enhanced Lead Conversion:** The analysis provides a robust framework for predicting lead conversion, potentially increasing the conversion rate significantly.

- **Focused Marketing Efforts:** By identifying key predictors of lead conversion, marketing strategies can be more targeted and efficient.

- **Continuous Improvement:** Regular updates and monitoring of the model will help maintain high performance and adapt to new data trends.

This comprehensive analysis empowers X Education to optimize its lead conversion processes, ultimately driving higher engagement and revenue from their marketing efforts.