Summary Report on Classification Model Performance

Introduction

This report summarizes the performance of a classification model aimed at predicting the likelihood of conversion based on various features. The model's predictions were evaluated using accuracy, confusion matrix, and other relevant metrics.

Model Predictions

- A new column, Predicted, was created in the training dataset (y_train_final) using a probability threshold of 0.5.
- The accuracy of the model at this threshold was **78.62%**.

Confusion Matrix

The confusion matrix for the predictions is as follows:

	Predicted 0	Predicted 1
Actual 0	2595	455
Actual 1	693	1627

From this confusion matrix:

- **True Negatives (TN)**: 2595
- **False Positives (FP)**: 455
- **False Negatives (FN)**: 693
- **True Positives (TP)**: 1627

Additional Metrics

- **Sensitivity (Recall)**: TP/(TP+FN) = 0.7013
- **Specificity**: TN/(TN+FP) = 0.8508
- **False Positive Rate (FPR)**: FP/(TN+FP) = 0.1492

ROC Curve Analysis

The ROC curve was plotted to visualize the trade-off between the true positive rate and false positive rate. The area under the curve (AUC) provided an aggregate measure of performance across all classification thresholds.

Optimal Cut-Off Threshold

Multiple probability cut-offs were tested to find an optimal threshold. The analysis suggested that a threshold of **0.4** yielded a balance between sensitivity and specificity, resulting in improved model performance.

Final Predictions

Using the optimal threshold of **0.4**, the final predictions were generated, and the confusion matrix was recalculated:

	Predicted 0	Predicted 1
Actual 0	2430	620
Actual 1	526	1794

Metrics at Threshold 0.4:

- **Accuracy**: 64.80%
- **Sensitivity**: 77.33%
- **Specificity**: 79.67%
- **Precision**: 74.32%
- **Recall**: 77.33%

Test Data Predictions

The model was applied to test data, and the final predictions were evaluated:

	Predicted 0	Predicted 1
Actual 0	1040	272
Actual 1	225	765

Metrics on Test Data

- **Accuracy**: 78.41%
- **Sensitivity**: 77.27%
- **Specificity**: 79.27%
- **Precision**: 73.77%
- **Recall**: 77.27%

Conclusion

The model demonstrates satisfactory performance with a balanced accuracy and other relevant metrics. Further tuning and exploration of features may enhance predictive capability.