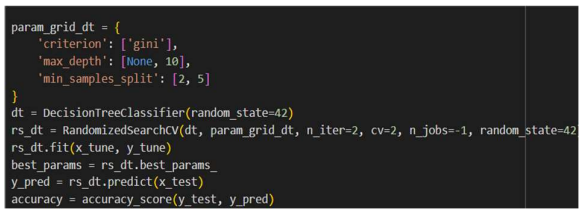
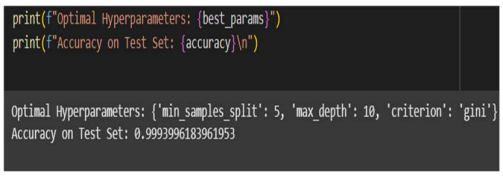
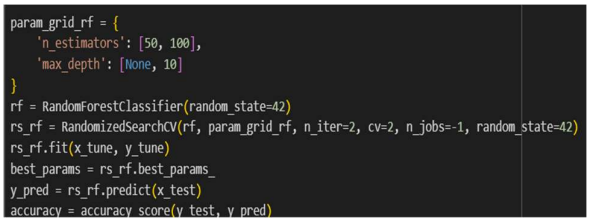
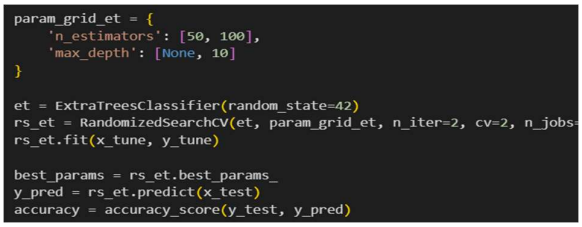
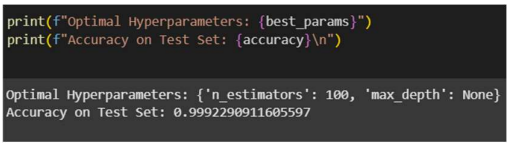
Model Optimization and Tuning Phase Report

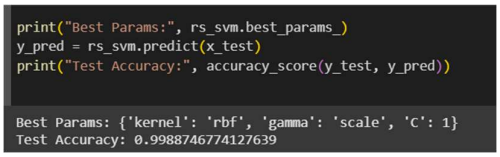
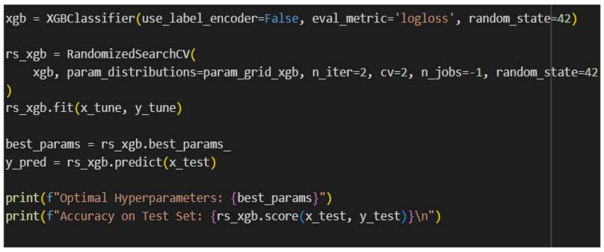
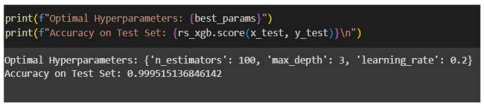
Date	15 June 2025
Team ID	SWTID1749662491
Project Title	Online Payments Fraud Detection using Machine Learning
Maximum Marks	10 Marks

Model Optimization and Tuning Phase

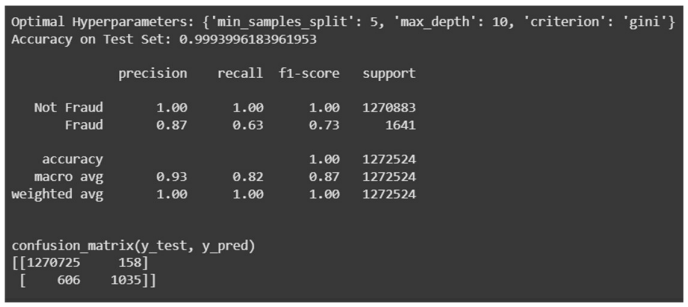
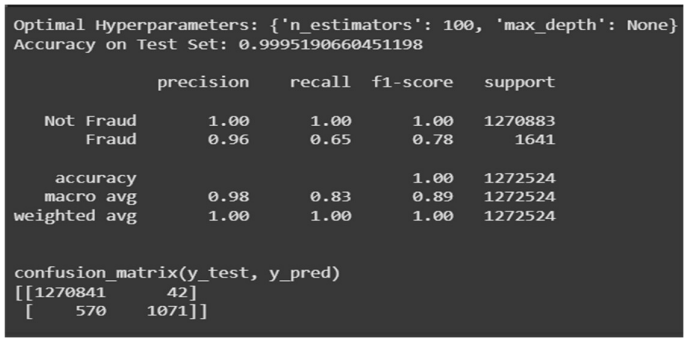
The Model Optimization and Tuning Phase involves refining machine learning models for peak performance. It includes optimized model code, fine-tuning hyperparameters, comparing performance metrics, and justifying the final model selection for enhanced predictive accuracy and efficiency.

Hyperparameter Tuning Documentation (6 Marks):

Model	Tuned Hyperparameters	Optimal values
Decision Tree		
Random Forest		
Extra Trees		

SVM		
XBG Classifier		

Performance Metrics Comparison Report (2 Marks):

Model	Optimized Metric
Decision Tree	
Random Forest	

Extra Trees	ExtraTrees Classifier Results precision recall f1-score support Not Fraud 1.00 1.00 1.00 1270883 Fraud 0.99 0.78 0.87 1641 accuracy 1.00 1272524 macro avg 0.99 0.89 0.93 1272524 weighted avg 1.00 1.00 1.00 1272524 confusion_matrix(y_test, y_pred) [[1270869 14] [369 1272]]
SVM	Support Vector Machine Classifier Results precision recall f1-score support Not Fraud 1.00 1.00 1.00 1270883 Fraud 0.99 0.21 0.35 1641 accuracy 1.00 1272524 macro avg 1.00 0.61 0.67 1272524 weighted avg 1.00 1.00 1.00 1272524 confusion_matrix(y_test, y_pred) [[1270880 3] [1295 346]]
XBG Classifier	Optimal Hyperparameters: {'n_estimators': 100, 'max_depth': 3, 'learning_rate': 0.2} Accuracy on Test Set: 0.999515136846142 precision recall f1-score support Not Fraud 1.00 1.00 1.00 1270883 Fraud 0.98 0.63 0.77 1641 accuracy 1.00 1272524 macro avg 0.99 0.82 0.89 1272524 weighted avg 1.00 1.00 1.00 1272524 confusion_matrix(y_test, y_pred) [[1270865 18] [599 1042]]

Final Model Selection Justification (2 Marks):

Final Model	**Reasoning**
XGBoost Classifier	The XGBoost model was selected as the final model due to its outstanding performance, achieving the highest accuracy and ROC-AUC scores during hyperparameter tuning and evaluation on the test set. XGBoost is well-known for its ability to efficiently handle large-scale datasets, capture complex feature interactions, and reduce overfitting through built-in regularization techniques. Its robustness, scalability, and superior predictive accuracy make it exceptionally well-suited for fraud detection tasks, aligning perfectly with the objectives of this project and justifying its selection as the final model.