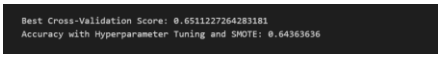
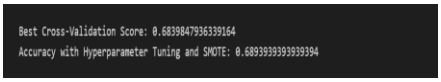
Model Optimization and Tuning Phase Template

Date	5 July 2024
Team ID	SWTID1720082658
Project Title	E-commerce Shipping Prediction 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.

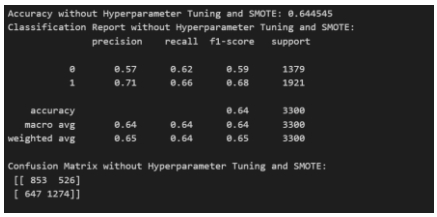
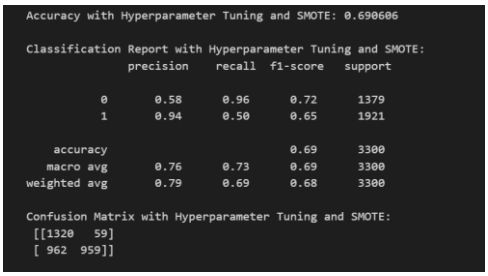
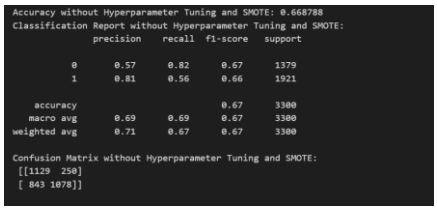
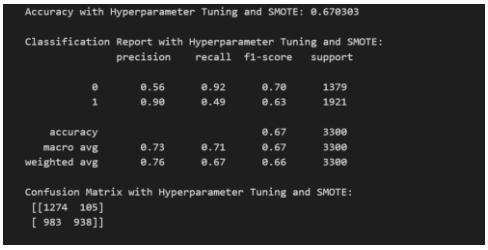
Hyperparameter Tuning Documentation (6 Marks):

Model	Tuned Hyperparameters	Optimal Values
LOGISTIC REGRESSION	Solver,penalty,C,max_iter	
RANDOM FOREST CLASSIFIER	n_estimators, max_depth, min_samples_split, min_samples_leaf	
KNN	n_neighbours,weights,metric,p	

XG BOOST	n_estimators,max_depth,learning_rate,subsample	Best Cross-Validation Score: 0.6822955536990625 Accuracy with Hyperparameter Tuning and SMOTE: 0.690606
SVC	Kernel,C,gamma	Best Cross-Validation Score: 0.6680090799389043 Accuracy with Hyperparameter Tuning and SMOTE: 0.670303

Performance Metrics Comparison Report (2 Marks):

Model	Baseline Metric	Optimized Metric
LOGISTIC REGRESSION	Accuracy without Hyperparameter Tuning and SMOTE: 0.6284848484848485 Classification Report without Hyperparameter Tuning and SMOTE: precision recall f1-score support 0 0.56 0.53 0.54 1379 1 0.67 0.70 0.69 1921 accuracy 0.63 3300 macro avg 0.62 0.61 0.61 3300 weighted avg 0.63 0.63 0.63 3300 Confusion Matrix without Hyperparameter Tuning and SMOTE: [[725 654] [572 1349]]	Accuracy with Hyperparameter Tuning and SMOTE: 0.64363636 Classification Report with Hyperparameter Tuning and SMOTE: precision recall f1-score support 0 0.55 0.76 0.64 1379 1 0.77 0.56 0.65 1921 accuracy 0.64 3300 macro avg 0.66 0.66 0.64 3300 weighted avg 0.68 0.64 0.64 3300 Confusion Matrix with Hyperparameter Tuning and SMOTE: [[1053 326] [850 1071]]
RANDOM FOREST CLASSIFIER	Accuracy without Hyperparameter Tuning and SMOTE: 0.690000 Classification Report without Hyperparameter Tuning and SMOTE: precision recall f1-score support 0 0.58 0.91 0.71 1379 1 0.89 0.53 0.67 1921 accuracy 0.69 3300 macro avg 0.74 0.72 0.69 3300 weighted avg 0.76 0.69 0.69 3300 Confusion Matrix without Hyperparameter Tuning and SMOTE: [[1250 129] [894 1027]]	Accuracy with Hyperparameter Tuning and SMOTE: 0.6893939393939394 Classification Report with Hyperparameter Tuning and SMOTE: precision recall f1-score support 0 0.58 0.95 0.72 1379 1 0.93 0.50 0.65 1921 accuracy 0.69 3300 macro avg 0.75 0.73 0.69 3300 weighted avg 0.78 0.69 0.68 3300 Confusion Matrix with Hyperparameter Tuning and SMOTE: [[1306 73] [952 969]]
KNN	Accuracy without Hyperparameter Tuning and SMOTE: 0.666969696969697 Classification Report without Hyperparameter Tuning and SMOTE: precision recall f1-score support 0 0.58 0.75 0.65 1379 1 0.77 0.61 0.68 1921 accuracy 0.67 3300 macro avg 0.68 0.68 0.67 3300 weighted avg 0.69 0.67 0.67 3300 Confusion Matrix without Hyperparameter Tuning and SMOTE: [[1035 344] [755 1166]]	Accuracy with Hyperparameter Tuning and SMOTE: 0.6633333333333333 Classification Report with Hyperparameter Tuning and SMOTE: precision recall f1-score support 0 0.57 0.81 0.67 1379 1 0.80 0.56 0.66 1921 accuracy 0.66 3300 macro avg 0.68 0.68 0.66 3300 weighted avg 0.70 0.66 0.66 3300 Confusion Matrix with Hyperparameter Tuning and SMOTE: [[1111 268] [843 1078]]

XG BOOST		
SVC		

Final Model Selection Justification (2 Marks):

Final Model	Reasoning
RANDOM FOREST CLASSIFIER	The RANDOM FOREST CLASSIFIER model was selected for its superior performance, exhibiting high accuracy during hyperparameter tuning. Its ability to handle complex relationships, minimize overfitting, and optimize predictive accuracy aligns with project objectives, justifying its selection as the final model