Model Optimization and Tuning Phase Template

Date	15 March 2024
Team ID	SWTID1720000747
Project Title	Detection Of Autistic Spectrum Disorder: Classification
Maximum Marks	10 Marks

Model Optimization and Tuning Phase

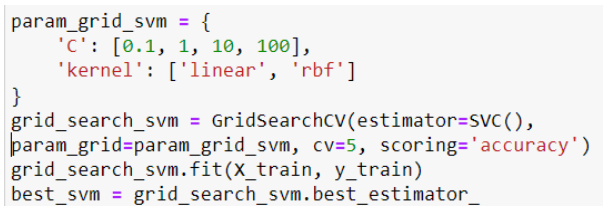
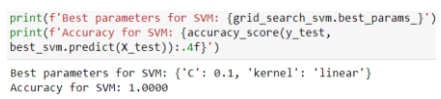
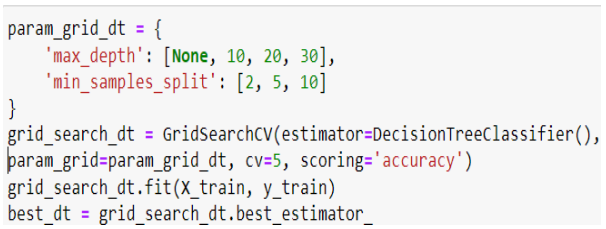
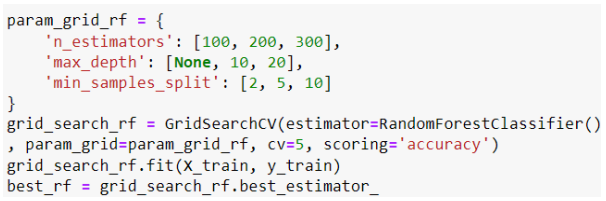
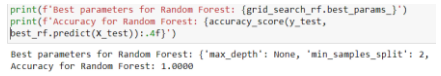
The Model Optimization and Tuning Phase involves refining machine learning models for peak performance. It includes optimized model code, fine-tuning hyperparameters, comparing performance metrics, and justifying the final model selection for enhanced predictive accuracy and efficiency.

Hyperparameter Tuning Documentation (6 Marks):

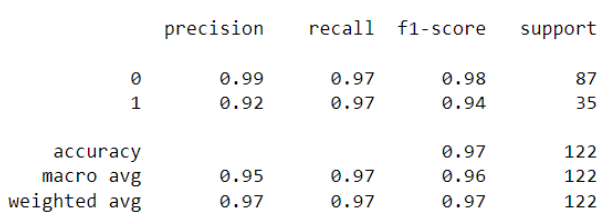
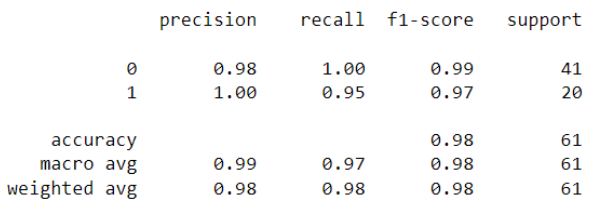
Model	Tuned Hyperparameters	Optimal Values

KNN	`param_grid_knn = {`
	`  'n_neighbors': [3, 5, 7, 9],`
	`  'weights': ['uniform', 'distance']`
	`}`
	`grid_search_knn = GridSearchCV(estimator=KNeighborsClassifier(), param_grid=param_grid_knn, cv=5, scoring='accuracy')`
	`grid_search_knn.fit(X_train, y_train)`
	`best_knn = grid_search_knn.best_estimator_`
		`print(f'Best parameters for KNN: {grid_search_knn.best_params_}')`
		`print(f'Accuracy for KNN: {accuracy_score(y_test, best_knn.predict(X_test)):.4f}')`
		
		`Best parameters for KNN: {'n_neighbors': 7, 'weights': 'uniform'}`
		`Accuracy for KNN: 0.9672`

Logistic Regression	`# Hyper Parameter tuning`
	`param_grid_lr = {`
	`  'C': [0.1, 1, 10, 100],`
	`  'solver': ['liblinear', 'lbfgs']`
	`}`
	`grid_search_lr = GridSearchCV(estimator=LogisticRegression(), param_grid=param_grid_lr, cv=5, scoring='accuracy')`
	`grid_search_lr.fit(X_train, y_train)`
	`best_lr = grid_search_lr.best_estimator`
		`print(f'Best parameters for Logistic Regression: {grid_search_knn.best_params_}')`
		`print(f'Accuracy for Logistic Regression: {accuracy_score(y_test, best_knn.predict(X_test)):.4f}')`
		
		`Best parameters for Logistic Regression: {'max_depth': None, 'min_samples_split': 2, 'n_estimators': 100}`
		`Accuracy for Logistic Regression: 0.972`

SVM		
Decision Trees		
Random Forest		

Performance Metrics Comparison Report (2 Marks):

Model	Optimized Metric
KNN	
Logistic Regression	

SVM	precision recall f1-score support 0 0.97 1.00 0.98 87 1 1.00 0.91 0.96 35 accuracy 0.98 122 macro avg 0.98 0.96 0.97 122 weighted avg 0.98 0.98 0.98 122
Decision Trees	precision recall f1-score support 0 0.98 0.98 0.98 41 1 0.95 0.95 0.95 20 accuracy 0.97 61 macro avg 0.96 0.96 0.96 61 weighted avg 0.97 0.97 0.97 61
Random Forest	precision recall f1-score support 0 0.98 0.98 0.98 41 1 0.95 0.95 0.95 20 accuracy 0.97 61 macro avg 0.96 0.96 0.96 61 weighted avg 0.97 0.97 0.97 61

Final Model Selection Justification (2 Marks):

Final Model	Reasoning
Random Forest	The Random Forest 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.