



Model Optimization and Tuning Phase Template

Date	10 July 2024
Team ID	SWTID1720097765
Project Title	xxxxxx
Maximum Marks	10 Marks

Model Optimization and Tuning Phase

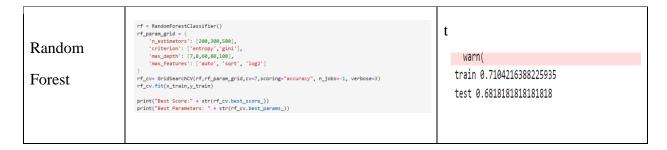
The Model Optimization and Tuning Phase involves refining machine learning models for peak performance. It includes optimized model code, fine-tuning hyperparameters, comparing performance metrics, and justifying the final model selection for enhanced predictive accuracy and efficiency.

Hyperparameter Tuning Documentation (6 Marks):

Model	Tuned Hyperparameters	Optimal Values
Logistic Regression	<pre>lg = LogisticRegressionCV(n_jobs=-1,random_state= 1234) lg_param_grid = { 'Cs': [6,8,10,15,20], 'max_iter': [60,80,100] } lg_cv= GridSearchCV(lg,lg_param_grid,cv=5,scoring="accuracy", n_jobs=-1, verbose=3) lg_cv.fit(x_train_normalized,y_train) print("Best Score:" + str(lg_cv.best_score_)) print("Best Parameters: " + str(lg_cv.best_params_))</pre>	Fitting 5 folds for each of 15 candidates, totalling 75 fits Best Score:0.6356404077730116 Best Parameters: {'Cs': 6, 'max_iter': 60}
SVM	<pre>svc = svm.SVC(random_state=1234,kernel='rbf',C= 10, gamma= 5 , tol = 1e-2,verbose = 1) svc.fit(x_train_normalized, y_train) print('train_score',svc.score(x_train_normalized,y_train)) print('test_score',svc.score(x_test_normalized,y_test))</pre>	[LibSVM]train score 0.6650755767700876 test score 0.6668181818181819
XG Boost	<pre>params = ('min_child_weight': (10,20), 'gmme: [1.5, 2.0, 2.5], 'colsample_bytree': [0.6, 0.8, 0.9], 'max_depth': [4,5,6] yp = NBC(lassifier(learning_rate-0.5, n_estimators-100, objective='binary:logistic', nthread-3) fitnodel: "of:dosen-cho(v(pg), param_grid-params, cvvs, refit=True, scoring='accuracy", n_jobs=-1, verbose=3) fitnodel: "fict_versin_promatize(s_j_versin) print(fitnodel.best_estimator_, fitnodel.best_params_, fitnodel.best_score_)</pre>	Sitting 5 min for man of the contineous, continuing 178 min statistical for a continuing continuing and continuing contin







Performance Metrics Comparison Report (2 Marks):

Model	Baseline Metric					Opt	imize	ed Metric			
	Name /	Accuracy	f1 score	Recall	Precision		Name	Accuracy	f1_score	Recall	Precision
MODELS	0 logistic regression	59.27	100000000	100.00	59.27	0	logistic regression	59.27	74.43	100.00	59.27
	1 logistic regression CV	63.27	68.78		71.99	1	logistic regression CV	63.27	66.78	62.27	71.9
	2 XGBoost	65.59	69.63	66.56	73.00	2	XGBoost	65.27	68.84	64.72	73.5
	3 Ridge classifier	59.27	74.43	100.00	59.27	3	Ridge classifier	59.27	74.43	100.00	59.2
	4 KNN	63.27	68.39	67.02	69.81	4	KNN	63.27	68.39	67.02	69.8
	5 Random Forest	67.18	70.38	65.72	75.71	5	Random Forest	67.55	66.38	54.06	85.9
	6 Support Vector Classifier	59.27	74.43	100.00	59.27	6	Support Vector Classifier	67.05	69.06	62.04	77.8

Final Model Selection Justification (2 Marks):

Final Model	Reasoning
	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
Random Forest	aligns with project objectives, justifying its selection as the final model