



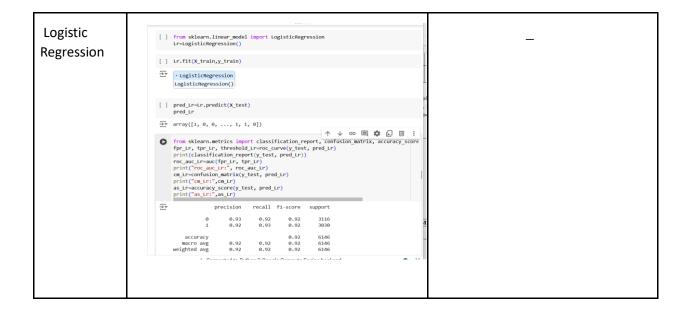
Model Optimization and Tuning Phase Report

Date	15 July 2024
Team ID	739930
Project Title	Airline Review 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):







KNN	[] from sklearn.neighbors import KNeighborsClassifier knn=KNeighborsClassifier(n_neighbors=5)
	[] knn.fit(X_train,y_train) - KNeighborsClassifier KNeighborsClassifier()
	pred_knn=knn.predict(X_test) pred_knn
	array([1, 0, 0,, 1, 1, 0]) [] from sklearn.metrics import classification report, confusion matrix, accuracy score
	<pre>fpr_knn, tpr_knn, threshold_knn = roc_curve(y_test, pred_knn) print(classification_report(y_test, pred_knn)) roc_auc_knn = auc(fpr_knn, tpr_knn) print('roc_auc_knn:', roc_auc_knn) cm_knn=confusion_matrix(y_test, pred_knn) print('cm_knn:',cm_knn) as_knn=accuracy_score(y_test, pred_knn) print("as_knn:",as_knn)</pre>
XGB	[] from xgboost import XGBClassifier xgb=XGBClassifier() [] xgb-fit(X_train,y_train)
	XGBClassifier (base score-None, booster-None, callbacks-None, colsample_bylevel-None, colsample_bylevel-None, colsample_bylevel-None, colsample_bytree-None, device-None, early_stopping_rounds-None, earble_categorical-False, eval_metri-c-None, feature_types-None, gamma-Hone, goognatical-False, eval_metri-c-None, feature_types-None, gamma-Hone, goognatical-False, eval_metri-c-None, importance_types-None, gamma-Hone, goognatical-False, calt_to_oneNone-hone, max_bin-None, max_cat_threshold-None, max_cat_to_oneNone, max_leaves-None, max_delta_step-Hone, max_depth-Hone, max_leaves-None, max_cat_tarts-None_max_leaves-None, max_tarts-goognatics-None, max_tarts-goognatics-None, max_tarts-goognatics-None, max_tarts-goognatics-None, max_tarts-goognatics-None, max_tarts-goognatics-None, max_tarts-goognatics-None, max_tarts-goognatics-None, moderatics-None, moderatics-None, moderatics-None,)
	[] pred_xgb-xgb.predict(X_test)
	from sklearn.metrics import classification_report, confusion_matrix, accuracy_score fpr_xgb, tpr_xgb, threshold_xgb-roc_curve (y_test, pred_xgb) print(classification_report(y_test, pred_xgb)) roc_auc_xgb-suc_xgb, rpc_xgb, pr_xgb) print("roc_auc_xgbs", roc_auc_xgb) cm_xgb-confusion_matrix(y_test, pred_xgb) print("cm_xgb;", cm_xgb) <pre>as_xgb-accuracy_score(y_test, pred_xgb) print("as_xgb;", as_xgb)</pre>

Performance Metrics Comparison Report (2 Marks):

Model	Optimized Metric						
Decision Tree	→		precision	recall	f1-score	support	
		0	0.95	0.95	0.95	3116	
		1	0.95	0.95	0.95	3030	
		accuracy			0.95	6146	
		macro avg	0.95	0.95	0.95	6146	
		weighted avg	0.95	0.95	0.95	6146	
	roc_auc_dt 0.9479143100446116 cm_dt: [[2958 158] [162 2868]] as_dt: 0.9479336153595834						





€	Y	precision	recal	l f1-score	support		
	0	0.95	0.96	6 0.96	3116		
		0.90	0.9.	0.90	3636		
	accuracy	,		0 96	61/16		
	-		0.00				
	•	,					
	weighted ave	0.96	0.90	0.96	6146		
	roc_auc_rfc 0.9582704194681343 cm_rfc: [[3003 113] [143 2887]] as_rfc: 0.9583468922876668						
		precision	recall	f1-score	support		
	_	2.25		0.05	2446		
	1	0.93	0.96	0.95	3030		
	200111201			0 OF	6146		
	_	0.05	0.05				
	_						
	weighted avg	0.95	0.95	0.95	6146		
	cm_knn: [[291 [133 2897]]	3 203]					
→	1	precision	recall	f1-score	support		
	0	0.96	0.97	0.96	3116		
	1	0.97	0.96	0.96	3030		
	accuracy			0.06	6146		
		0 06	0 06				
	roc_auc_xgb: 0 cm_xgb: [[3011 [122 2908]]	.96301946305 105]	02845				
		accuracy macro avg weighted avg roc_auc_rfc cm_rfc: [[36	0 0.95 1 0.96 accuracy macro avg 0.96 weighted avg 0.96 roc_auc_rfc 0.9582704194 cm_rfc: [[3003 113] [143 2887]] as_rfc: 0.95834689228766 precision 0 0.96 1 0.93 accuracy macro avg 0.95 weighted avg 0.95 roc_auc_knn: 0.9454789927 cm_knn: [[2913 203] [133 2897]] as_knn: 0.945330296127562 precision 0 0.96 1 0.97 accuracy macro avg 0.96 veighted avg 0.96 roc_auc_xgb: 0.96301946305 cm_xgb: [[3011 105] [122 2908]]	0 0.95 0.96 1 0.96 0.95 accuracy macro avg 0.96 0.96 weighted avg 0.96 0.96 roc_auc_rfc 0.9582704194681343 cm_rfc: [[3003 113] [143 2887]] as_rfc: 0.9583468922876668 → precision recall 0 0.96 0.93 1 0.93 0.96 accuracy macro avg 0.95 0.95 weighted avg 0.95 0.95 roc_auc_knn: 0.945478992700297 cm_knn: [[2913 203] [133 2897]] as_knn: 0.9453302961275627 → precision recall 0 0.96 0.97 1 0.97 0.96 accuracy macro avg 0.96 0.96 weighted avg 0.96 0.96 roc_auc_xgb: 0.9630194630502845 cm_xgb: [[3011 105]	0 0.95 0.96 0.96 1 0.96 0.95 0.96 1 0.96 0.95 0.96 accuracy 0.96 0.96 0.96 weighted avg 0.96 0.96 0.96 veighted avg 0.96 0.96 0.96 roc_auc_rfc 0.9582704194681343 cm_rfc: [[3003 113] [143 2887]] as_rfc: 0.9583468922876668 → precision recall f1-score 0 0.96 0.93 0.95 1 0.93 0.96 0.95 accuracy 0.95 0.95 0.95 weighted avg 0.95 0.95 0.95 roc_auc_knn: 0.945478992700297 cm_knn: [[2913 203] [133 2897]] as_knn: 0.9453302961275627 → precision recall f1-score 0 0.96 0.97 0.96 1 0.97 0.96 0.96 accuracy 0.96 0.96 0.96 weighted avg 0.96 0.96 0.96 weighted avg 0.96 0.96 0.96 veighted avg 0.96 0.96 0.96 roc_auc_xgb: 0.9630194630502845 cm_xgb: [[3011 105] [122 2998]]	0 0.95 0.96 0.96 3116 1 0.96 0.95 0.96 3030 accuracy 0.96 6146 macro avg 0.96 0.96 0.96 6146 weighted avg 0.96 0.96 0.96 0.96 6146 roc_auc_rfc 0.9582704194681343 cm_rfc: [[3003 113] [143 2887]] as_rfc: 0.9583468922876668 → precision recall f1-score support 0 0.96 0.93 0.95 3116 1 0.93 0.96 0.95 3030 accuracy 0.95 0.95 0.95 6146 weighted avg 0.95 0.95 0.95 6146 weighted avg 0.95 0.95 0.95 6146 roc_auc_knn: 0.945478992700297 cm_knn: [[2913 203] [133 2897]] as_knn: 0.9453302961275627 → precision recall f1-score support 0 0.96 0.97 0.96 3116 1 0.97 0.96 0.96 3030 accuracy 0.96 0.96 0.96 6146 weighted avg 0.96 0.96 0.96 6146 roc_auc_xgb: 0.9630194630502845 cm_xgb: [[3011 105] [122 2908]]	0 0.95 0.96 0.96 3116 1 0.96 0.95 0.96 3030 accuracy 0.96 0.96 6146 macro avg 0.96 0.96 0.96 6146 weighted avg 0.96 0.96 0.96 6146 roc_auc_rfc 0.9582704194681343 cm_rfc: []3003 113] []143 2887]] as_rfc: 0.9583468922876668 → precision recall f1-score support 0 0.96 0.93 0.95 3116 1 0.93 0.96 0.95 3030 accuracy 0.95 0.95 0.95 6146 weighted avg 0.95 0.95 0.95 6146 weighted avg 0.95 0.95 0.95 6146 roc_auc_knn: 0.945478992700297 cm_knn: [[2913 203] []133 2897]] as_knn: 0.9453302961275627 → precision recall f1-score support 0 0.96 0.97 0.96 3116 1 0.97 0.96 0.96 3030 accuracy 0.96 0.96 0.96 6146 weighted avg 0.96 0.96 0.96 6146 roc_auc_xgb: 0.9630194630502845 cm_xgb: [[3011 105] []122 2908]]





Final Model Selection Justification (2 Marks):

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