Project Development Phase
Model Performance Test

Date	11 November 2022
Team ID	PNT2022TMID36161
Project Name	Project - Developing a Flight Delay Prediction Model using Machine Learning
Maximum Marks	10 Marks

Model Performance Testing:

Project team shall fill the following information in model performance testing template.

S.No.	Parameter	Values	Screenshot

1.	Metrics	**Classification Model:** Confusion Matrix , Accuracy Score & Classification Report	In [24]:
			`#Model Evaluation`
			`from sklearn.metrics import accuracy_score,confusion_matrix, classification_report`
			`print(accuracy_score(y_test, pred))`
			0.9163899788711138
			
			In [26]:
			`print(confusion_matrix(y_test, pred))`
			[[2732 164]
			[113 304]]
			
			In [27]:
			`print(classification_report(y_test, pred))`
			precision recall f1-score support
			
			0.0 0.96 0.94 0.95 2896
			1.0 0.65 0.73 0.69 417
			
			accuracy 0.92 3313
			macro avg 0.80 0.84 0.82 3313
			weighted avg 0.92 0.92 0.92 3313

2.	Tune the Model	Hyperparameter Tuning , Validation Method	In [31]:
			`from sklearn.model_selection import cross_val_score, KFold, GridSearchCV`
			`kf = KFold(n_splits = 6, shuffle = True, random_state = 25)`
			`params = {'max_depth': [4,5,6],`
			`'min_samples_split': [2,3,4],`
			`'criterion': ['gini', 'entropy', 'log_loss']}`
			
			In [32]:
			`grid_cv = GridSearchCV(clf, params, cv = kf)`
			`grid_cv.fit(X_train,y_train)`
			
			Out[32]:
			GridSearchCV
			estimator: DecisionTreeClassifier
			DecisionTreeClassifier
			
			In [35]:
			`grid_cv.best_params_`
			Out[35]: {'criterion': 'entropy', 'max_depth': 6, 'min_samples_split': 2}
			
			In [34]:
			`cv_results = cross_val_score(clf, X_train,y_train, cv = kf)`
			`print(cv_results)`
			[0.92552366 0.9193173 0.91925466 0.92934783 0.92313665 0.9060559]