Model Performance Test

Date	10 November 2022
Team ID	PNT2022TMID16214
Project Name	Efficient Water Quality Analysis and Prediction using Machine Learning
Maximum Marks	10 Marks

Model Performance Testing:

S.No.	Parameter	Values	Screenshot

1.	Metrics	**Regression Model:**
		MAE - , MSE - , RMSE - , R2 score -
			Model Evaluation
			In [37]:
			`from sklearn import metrics`
			`print('MAE:',metrics.mean_absolute_error(y_test,y_pred))`
			`print('MSE:',metrics.mean_squared_error(y_test,y_pred))`
			`print('RMSE:',np.sqrt(metrics.mean_squared_error(y_test,y_pred)))`
			MAE: 0.4550025062656734
			MSE: 2.5859671077694255
			RMSE: 1.6080942471663238
			In [38]:
			`metrics.r2_score(y_test, y_pred)`
			Out[38]: 0.9759652869193766

2.	Tune the Model	Hyperparameter Tuning - Validation Method -	**Hyperparameter Tuning** In [ ]: ``` from sklearn.model_selection import cross_val_score, GridSearchCV ``` In [ ]: ``` param_grid = { 'bootstrap': [True], 'max_depth': [5, 10, None], 'max_features': ['auto', 'log2'], 'n_estimators': [5, 6, 7, 8, 9 ``` In [ ]: ``` rfr = RandomForestRegressor(random_state = 1) g_search = GridSearchCV(estimator = rfr, param_grid = param_grid, cv = 3, n_jobs = 1, verbose = 0, return_train_score=True) ``` In [ ]: ``` g_search.fit(x_train, y_train) print(g_search.best_params_) ``` ``` {'bootstrap': True, 'max_depth': 10, 'max_features': 'auto', 'n_estimators': 15} ``` ***Validation Method Cross validation*** In [ ]: ``` scores = cross_val_score(regressor, y_test, y_pred, cv=10, scoring='neg_mean_absolute_error') print(scores) ``` ``` [-0.88937508 -0.2277642 -0.62957576 -0.28678912 -0.52877112 -0.33818409 -0.59450265 -0.16186615 -0.17046191 -1.16749981] ```