**Project Development Phase**
**Model Performance Test**

| | |
|---|---|
| Date | 10 November 2022 |
| Team ID | PNT2022TMID21972 |
| Project Name | Project – Efficient Water Quality Analysis and Prediction 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 | **Regression Model:**<br>MAE - , MSE - , RMSE - , R2 score - | **Model Evaluation**<br>In [37]:<br>`from sklearn import metrics`<br>`print('MAE:',metrics.mean_absolute_error(y_test,y_pred))`<br>`print('MSE:',metrics.mean_squared_error(y_test,y_pred))`<br>`print('RMSE:',np.sqrt(metrics.mean_squared_error(y_test,y_pred)))`<br>MAE: 0.4550025062656734<br>MSE: 2.5859671077694255<br>RMSE: 1.6080942471663238<br>In [38]:<br>`metrics.r2_score(y_test, y_pred)`<br>Out[38]: 0.9759652869193766 |

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