Project Development Phase
Model Performance Test

Date	10 NOvember 2022
Team ID	PNT2022TMID21577
Project Name	Project - Car Resale Value Prediction
Maximum Marks	10 Marks

Model Performance Testing:

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

S.No.	Parameter	Values	Screenshot

1.	Metrics	**Regression Model:** MAE - , MSE - , RMSE - , R2 score -	In [25]:
			`from sklearn.metrics import mean_squared_error,mean_absolute_error`
			`mse = mean_squared_error(Y_test, y_pred)`
			`print(mse)`
			`11837192.971239958`
			
			In [26]:
			`rmse = np.sqrt(mse)`
			`print(rmse)`
			`mae = mean_absolute_error(Y_test, y_pred)`
			`print(mae)`
			`3440.5221945570934`
			`1635.1608915188156`
			
			In [17]:
			`y_pred = regressor.predict(X_test)`
			`print(r2_score(Y_test,y_pred))`
			`0.834527626497731`

2.	Tune the Model	Hyperparameter Tuning - **n_estimators** = [5,20,50,100] **max_features** = ['auto', 'sqrt'] **max_depth** = [10-120] **min_samples_split** = [2, 6, 10] **min_samples_leaf** = [1, 3, 4] **bootstrap** = [True, False]
		
		Validation Method - **RandomisedGridSearchCV**
			In [33]:
			`n_estimators = [5,20,50,100]`
			`max_features = ['auto', 'sqrt']`
			`max_depth = [int(x) for x in np.linspace(10, 120, num = 12)]`
			`min_samples_split = [2, 6, 10]`
			`min_samples_leaf = [1, 3, 4]`
			`bootstrap = [True, False]`
			
			`random_grid = {'n_estimators': n_estimators,`
			
			`'max_features': max_features,`
			
			`'max_depth': max_depth,`
			
			`'min_samples_split': min_samples_split,`
			
			`'min_samples_leaf': min_samples_leaf,`
			
			`'bootstrap': bootstrap}`
			
			In [34]:
			`from sklearn.model_selection import RandomizedSearchCV`
			`rf_random = RandomizedSearchCV(estimator = regressor,param_distributions = random_grid,`
			`n_iter = 100, cv = 5, verbose=2, random_state=35, n_jobs = -1)`
			
			In [36]:
			`rf_random.fit(X_train, Y_train)`
			
			`Fitting 5 folds for each of 100 candidates, totalling 500 fits`
			
			`C:\ProgramData\Anaconda3\lib\site-packages\sklearn\model_selection\_search.py:926: DataC`
			`passed when a 1d array was expected. Please change the shape of y to (n_samples,), for ex`
			`self.best_estimator_.fit(X, y, **fit_params)`
			
			Out[36]:
			`RandomizedSearchCV(cv=5,`
			`estimator=RandomForestRegressor(max_depth=10,`
			`n_estimators=1000,`
			`random_state=34),`
			`n_iter=100, n_jobs=-1,`
			`param_distributions={'bootstrap': [True, False],`
			`'max_depth': [10, 20, 30, 40, 50, 60,`
			`70, 80, 90, 100, 110,`
			`120],`
			`'max_features': ['auto', 'sqrt'],`
			`'min_samples_leaf': [1, 3, 4],`
			`'min_samples_split': [2, 6, 10],`
			`'n_estimators': [5, 20, 50, 100]},`
			`random_state=35, verbose=2)`