

Project Development Phase Model Performance Test

Date	10 November 2023
Team ID	592766
Project Name	Walmart Sales Analysis for Retail Industry 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	Random Forest: MAE – 1644.44 MSE – 16414124.70 RMSE – 4051.43 R2 score – 0.96 Accuracy – 96.84 XgBoost: MAE – 3068.66 MSE – 30796191.59 RMSE – 5549.43 R2 score – 0.94 Accuracy – 94.08	 <pre> from sklearn.ensemble import RandomForestRegressor rf = RandomForestRegressor (n_estimators=50, max_depth=20, min_samples_split=3, rf.fit(X_train, y_train) print ('Accuracy:',rf.score(X_test, y_test)*100, '%') y_pred = rf.predict(X_test) <ipython-input-74-ca2511d819d8>:3: DataConversionWarning: A column-vector y was rf.fit(X_train, y_train) Accuracy: 96.8495179584525 % print('MSE: ', mean_squared_error(y_test, y_pred, squared=True)) print('RMSE: ', mean_squared_error(y_test, y_pred, squared=False)) print('MAE: ', mean_absolute_error (y_test, y_pred)) print('R2: ', explained_variance_score (y_test, y_pred)) MSE: 16414124.70616664 RMSE: 4051.434894721454 MAE: 1644.4428512985985 R2: 0.9684952121274976 </pre>  <pre> pred = xg_reg.predict(X_train) y_pred = xg_reg.predict(X_test) print ('Accuracy: ',xg_reg.score(X_test, y_test)*100, '%') Accuracy: 94.08906358198728 % print('MSE: ', mean_squared_error(y_test, y_pred, squared=True)) print('RMSE: ', mean_squared_error(y_test, y_pred, squared=False)) print('MAE: ', mean_absolute_error (y_test, y_pred)) print('R2: ', explained_variance_score (y_test, y_pred)) MSE: 30796191.593139805 RMSE: 5549.43164595617 MAE: 3068.662971047117 R2: 0.9408906894277229 </pre>
2.	Tune the Model	Validation Method - cross validation	Random Forest:  <pre> cv = cross_val_score(rf,X,y,cv=6) np.mean(cv) 0.7280028435919697 </pre> XgBoost:  <pre> cv = cross_val_score(xg_reg,X,y,cv=6) np.mean(cv) 0.7482499257941506 </pre>