REPORT

Predicting House Prices:

An End-to-End Regression Modelling Workflow

Dataset: (14620, 23)

Baseline metrics:

mean:- RMSE: 371483.11385227746, MAE: 238078.25671654465, R2: -0.0005092836256821442, MAPE: 53.59734014997075

median:- RMSE: 383555.6556708793, MAE: 227152.43593251254, R2: -0.06659550775878964, MAPE: 42.95768455496701

group_mean_by_number of bedrooms:- RMSE: 349935.0734978826, MAE: 221090.67183646534, R2: 0.11219432852611488, MAPE: 48.33849225997894

Using strong predictor: living area

SLR metrics: RMSE: 259895.68017691662,

MAE: 173534.2565043719,

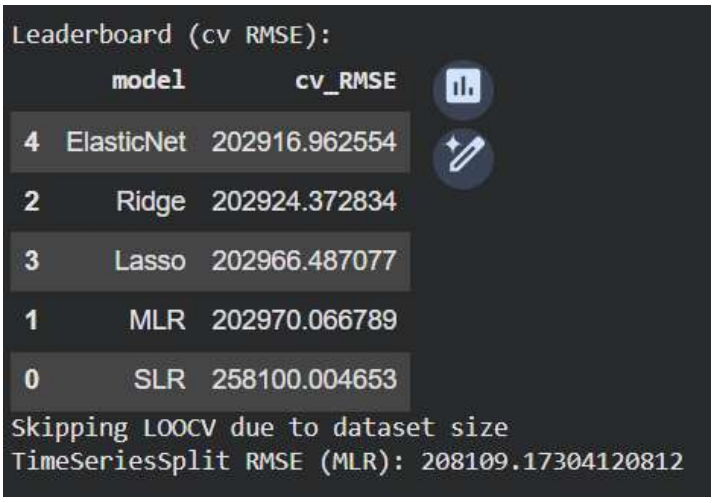
R2: 0.5102876120187136,

MAPE: 35.33774654032697

Ridge best params: regressor_alpha: 50} | val RMSE: 195318.25

Lasso best params: regressor_alpha: 50} | val RMSE: 195345.48

ElasticNet best params: regressor_alpha: 0.1 | val RMSE: 195398.11



Champion: Ridge with val RMSE: 195318.25

Test metrics: RMSE: 197809.02432231724, MAE: 125082.31170389932, R2: 0.7016842856620028, MAPE: 24.973364864080427