Model Evaluation

To determine the best machine learning model and the appropriate evaluation metrics for our data, we need to consider several factors:

First of all,

1: - Problem Type: -

We need to understand the type of the problem so our target feature is Price which is continuous so the problem type is REGRESSION.

2: - Model Selection: -

For Regression we can use the Linear Regression , Ridge regression, LASSO, Decision trees, Xg-boost and Random forests.

3: - Evaluation Metrics: -

For Regression we can use metrics like Root mean squared error(RMSE), Mean absolute error(MAE), Adjusted R-squared and R-squared .

Mean Absolute error (MAE):-

Provides an average of the absolute errors between predicted and actual values in regression model. It is calculated by taking the average of the absolute differences between the predicted and actual values for each data point. MAE is less sensitive to outliers compared to MSE.

Root Mean squared error (RMSE): -

Similar to MAE but gives higher weight to larger errors. RMSE is a popular metric for evaluating the accuracy of a model's predictions. RMSE is often preferred for interpretation because it provides a more intuitive measure of error. And also is easier to interpret.

R-squared: -

R- squared represents the variance in the dependent variable that is predictable from the independent variables in regression model. It ranges from 0 to 1 , where 0 indicates model doesn’t explain any variability in the target variable, and 1 indicates that the model perfectly explain the variability.

Adjusted R-squared: -

Adjusted R-squared ranges from negative infinity to 1. A higher value indicates a better fit of the model, with 1 it represents a perfect fit. Adjusted R-squared penalizes overcomplication, which helps avoid overfitting and offers a more cautious evaluation of a regression model's goodness of fit.