Lab4_PolyRegression_model

This homework assignment includes two models, linear regression model using a single feature and polynomial regression model. Both the model accepts the single input features and predict house's price, using StandardScaler.

a) Which model gives you better performance: Multiple Linear Regression or Polynomial Regression?

Linear Regression gives better performance in this case, because it has a significantly higher R^2 score (0.42), showing that it explains more variance in the target variable.

Although the polynomial model has a slightly lower MSE, its R^2 score of 0.0 indicates it's not capturing any meaningful relationship.

Thus, linear regression fits the dataset more appropriately for this problem, it explains the relationship between square footage and house price better than polynomial models.

b) How do you decide the optimal degree for Polynomial Regression?

To find the optimal polynomial degree, I tested polynomial regression models with increasing degrees from **1 to 6**.

For each degree, I evaluated key performance metrics:

R^2 Score (explained variance): Higher values mean better model fit.

Mean Squared Error (MSE) and **Mean Absolute Error (MAE)**: Lower values indicate better prediction accuracy.

I focused on:

1. **R^2 Score:** To ensure the model explains the target variable well.
2. **MSE/MAE:** To measure error and prediction accuracy.
3. **Model Complexity:** To avoid unnecessary complexity.

In this case, increasing the degree beyond 1 didn't improve the model's ability to explain variance (R^2 stayed at 0), and the slight reductions in MSE/MAE were not enough to justify the additional complexity.

Too Low Degree : Model becomes **too simple** (underfitting), missing important patterns. Predictions will have high error and low accuracy.

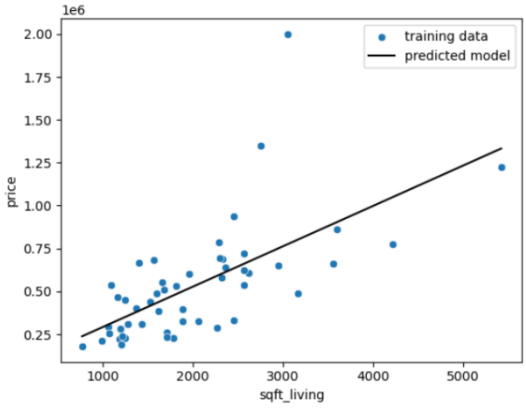
Too High Degree : Model becomes **too complex** (overfitting), fitting noise rather than the true relationship. This results in poor performance on new/unseen data and unstable predictions.

Linear Regression Model Result

r2-score: 0.42

mean squared error: 62185849719.22

mean absolute error: 161171.37



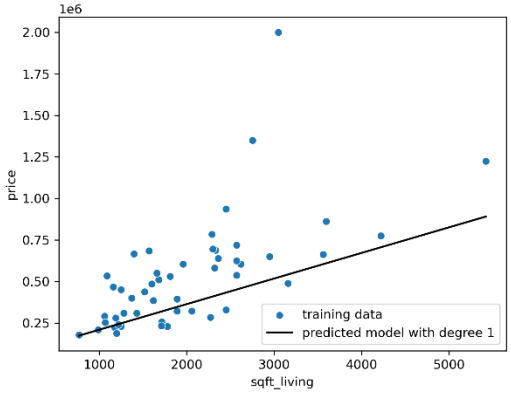
Polynomial Regression Model Result

Degree 1

r2-score: 0.0

mean squared error: 62185849719.0

mean absolute error: 161171.0

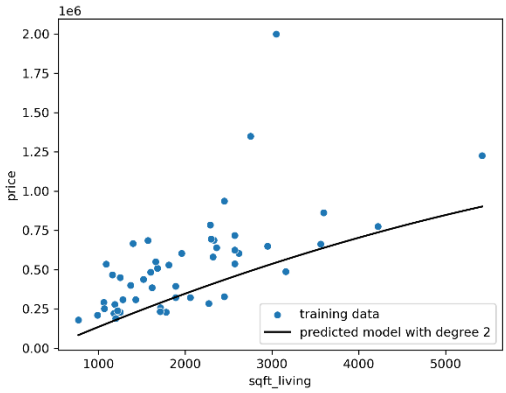


Degree 2

r2-score: 0.0

mean squared error: 60999286388.0

mean absolute error: 157625.0

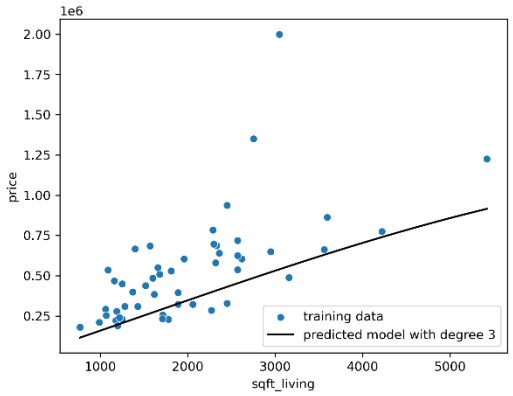


Degree 3

r2-score: 0.0

mean squared error: 60956300634.0

mean absolute error: 159361.0

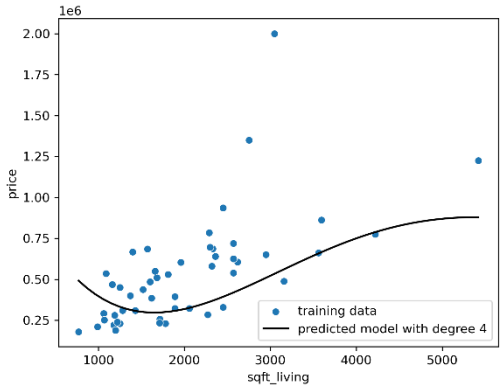


Degree 4

r2-score: 0.0

mean squared error: 58826504907.0

mean absolute error: 159294.0

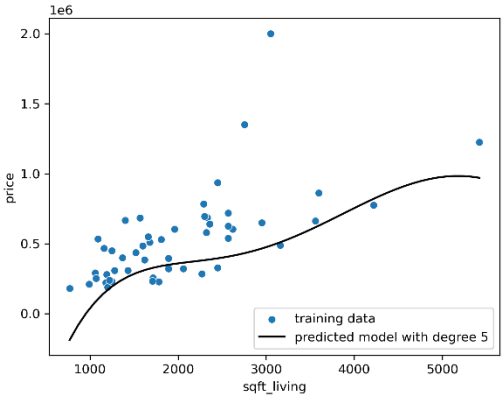


Degree 5

r2-score: 0.0

mean squared error: 55715230172.0

mean absolute error: 161830.0



Degree 6

r2-score: 0.0

mean squared error: 55042647383.0

mean absolute error: 162857.0