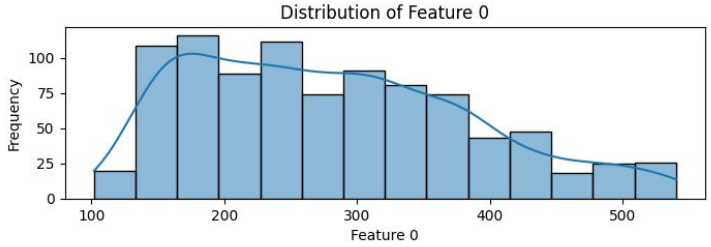
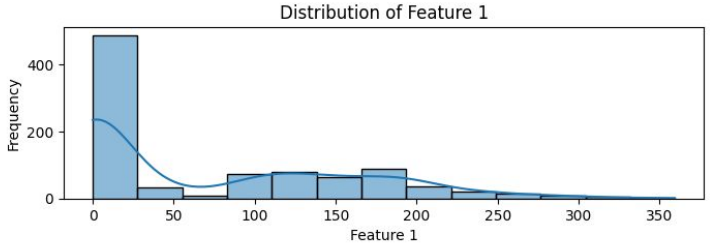
Quantitative Foundations

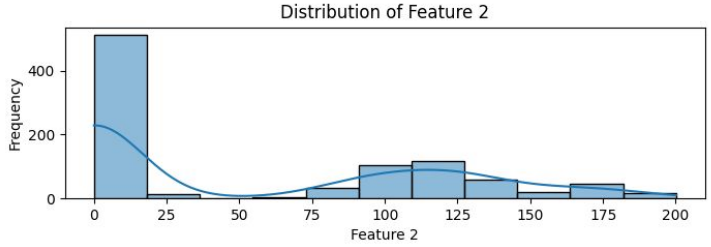
Linear Feature Engineering

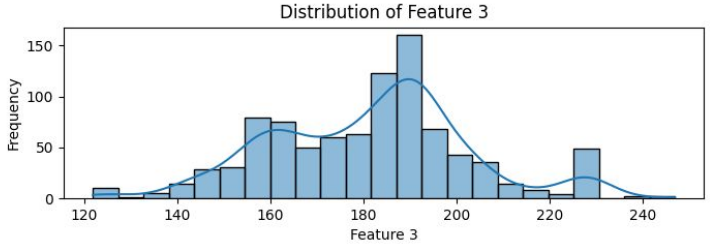
Image Adhikari and Suraj Poudel

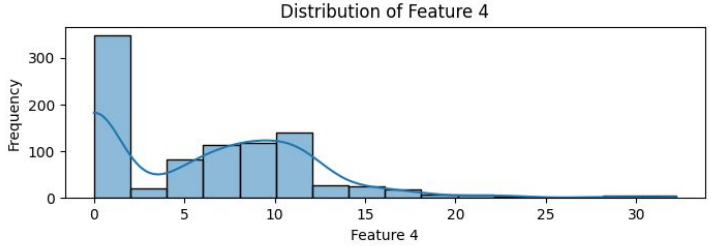
Data Distribution

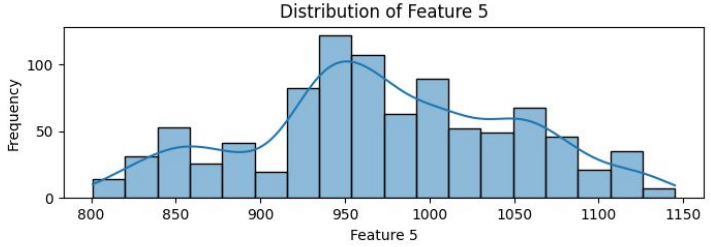


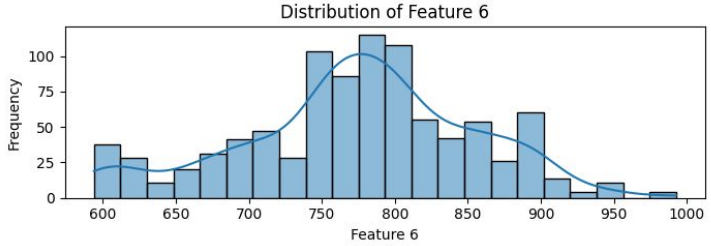


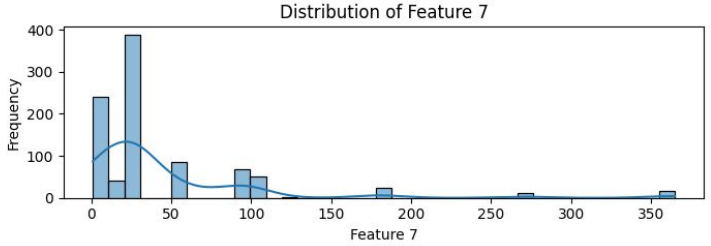




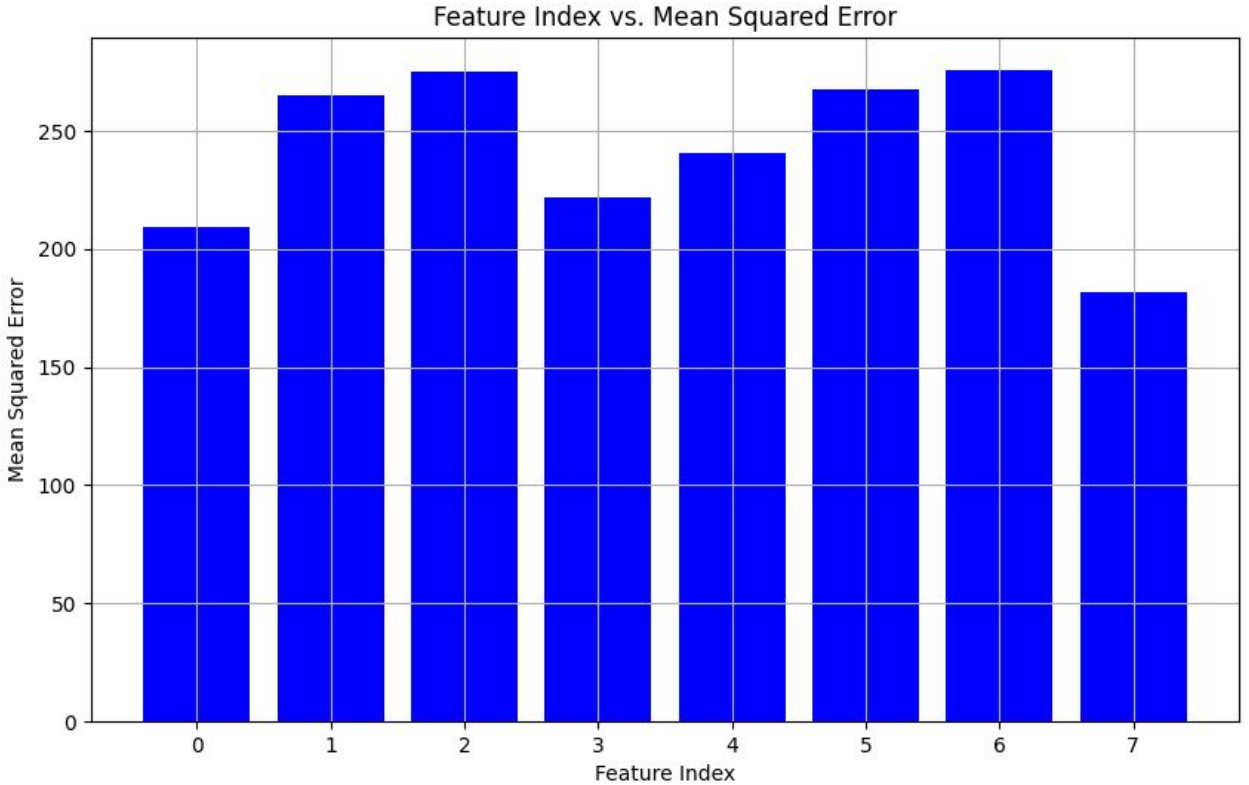


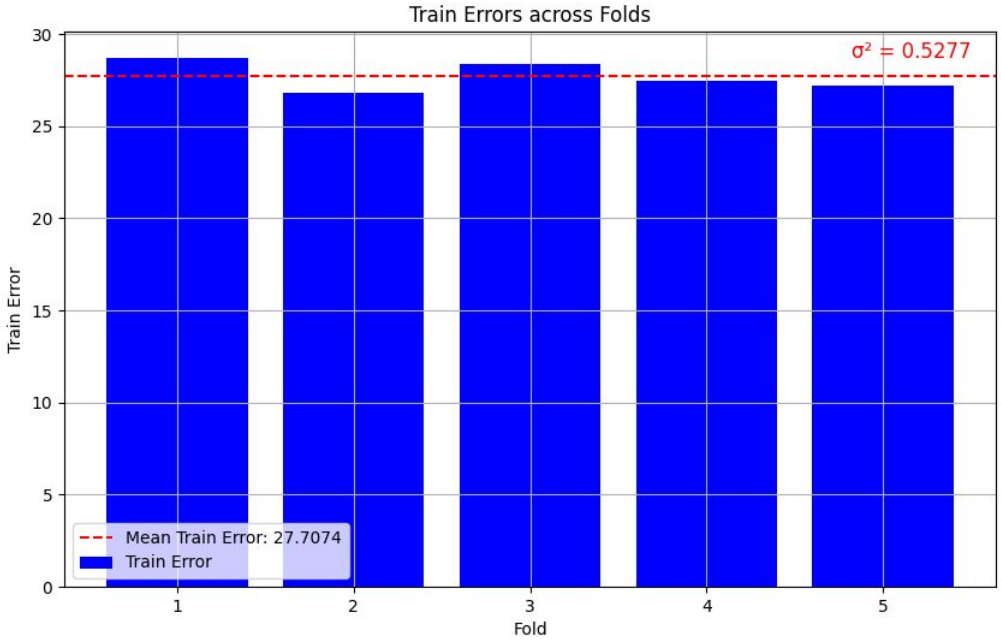




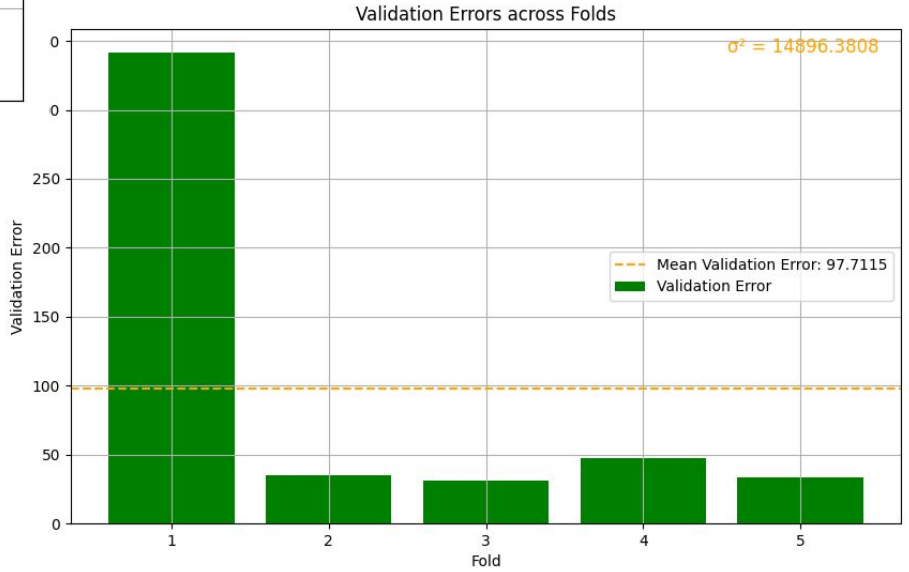


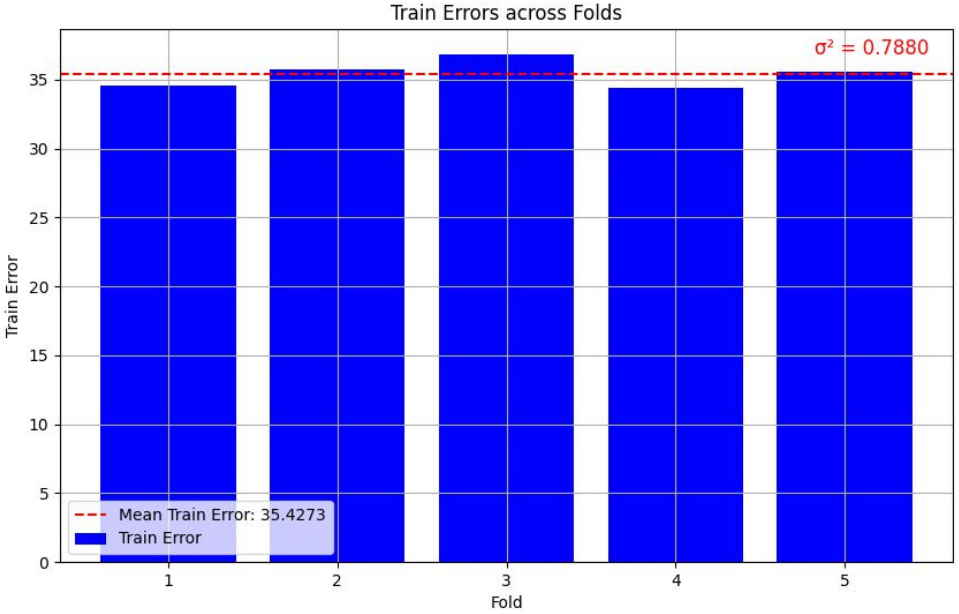
Feature Selection



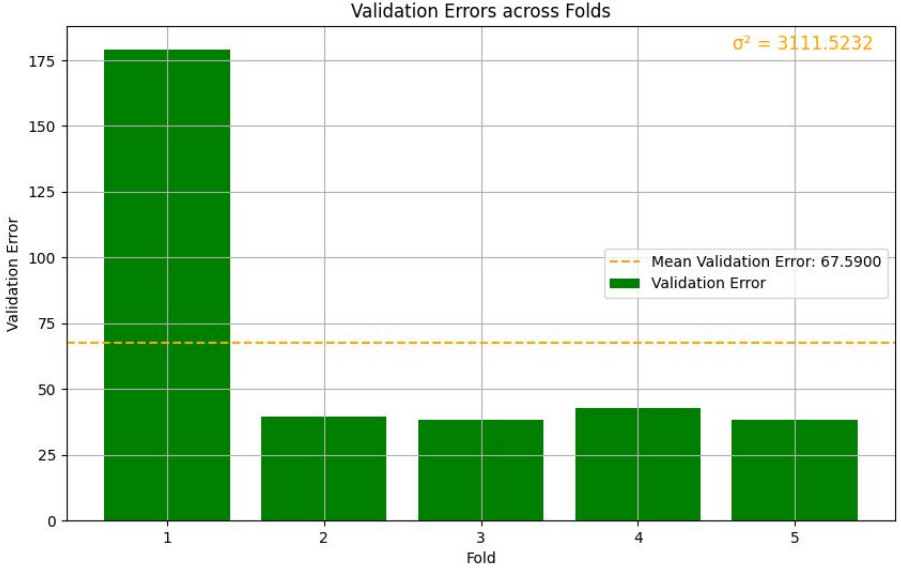


Feature Selection





Feature Selection



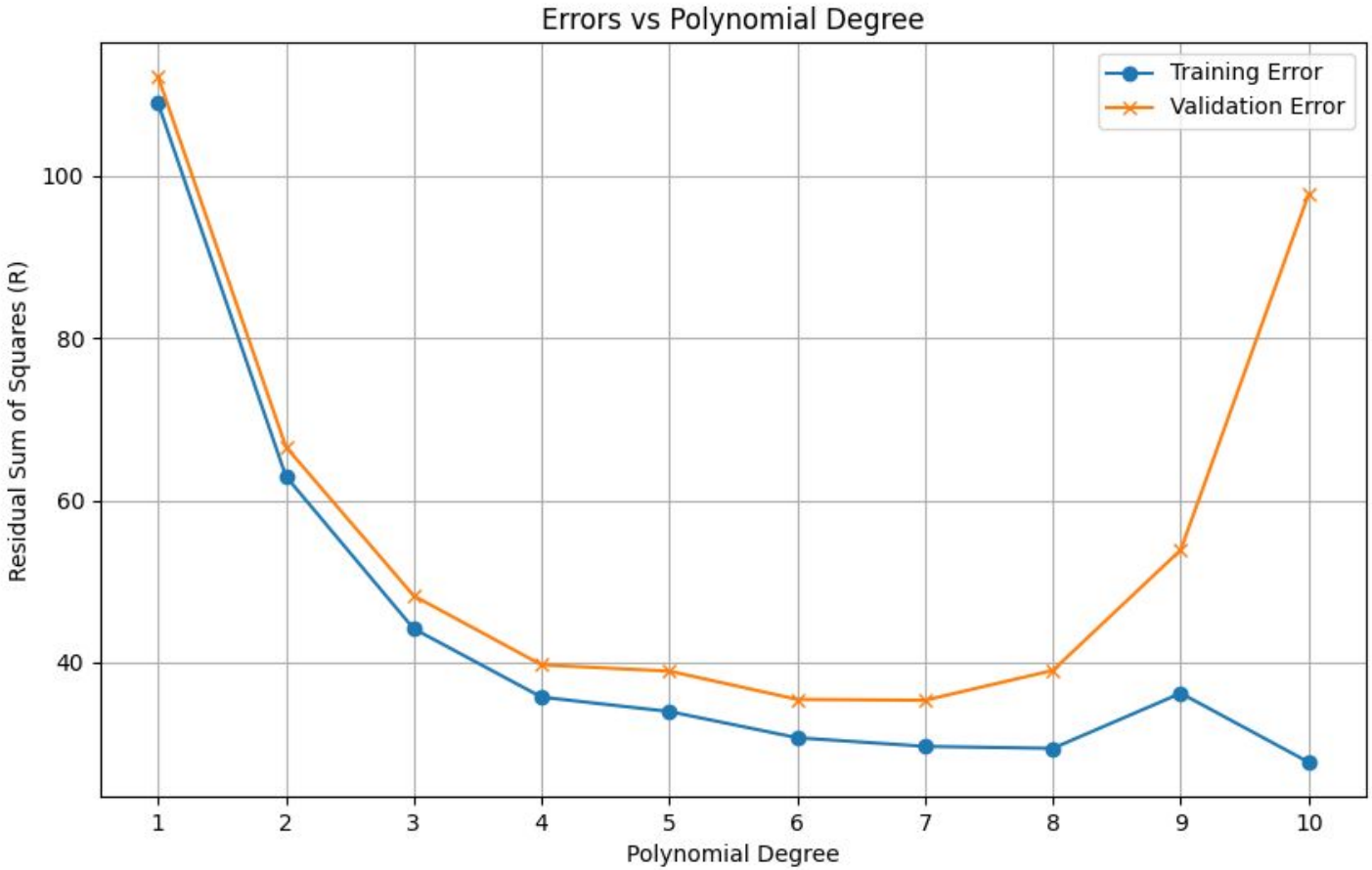
Fitting the model

Polynomial expansion

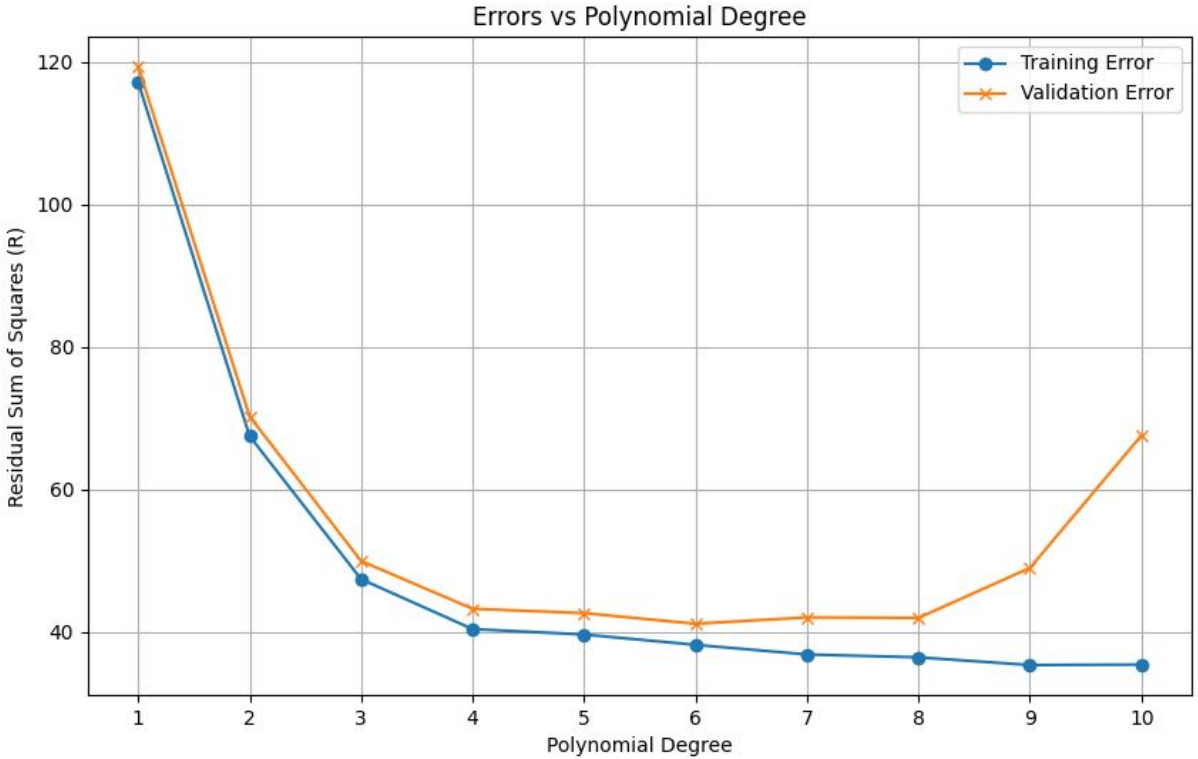
$$Given\ x_i,\ generate\ \{x_i^{\,p}\}_{p=1}^{\,d}$$

$$where\ d = 10$$

Search for best degree of polynomial



Search for best degree of polynomial



Prediction of test error

- **Estimated Test Error:** The mean squared test error is predicted to be 67.59.
- **Validation Process:** This value was obtained through K-fold cross-validation.
- **Feature Selection:** Optimal model features were selected to achieve this validation error.
- **Test Error Approximation:** We believe the validation error serves as a reliable approximation of the mean squared error on the test set.

Overfitting Mitigation

1. Cross Validation

2. Feature Selection

Thank You!