

Training and Test Splits

Splitting your data into a training and a test set can help you choose a model that has better chances at generalizing and is not overfitted.

The training data is used to fit the model, while the test data is used to measure error and performance.

Training error tends to decrease with a more complex model.

Polynomial Regression

Polynomial terms help you capture nonlinear effects of your features.

Other algorithms that help you extend your linear models are:

- Logistic Regression
- K-Nearest Neighbors
- Decision Trees
- Support Vector Machines
- Random Forests
- Ensemble Methods
- Deep Learning Approaches