Underfitting and Overfitting 101

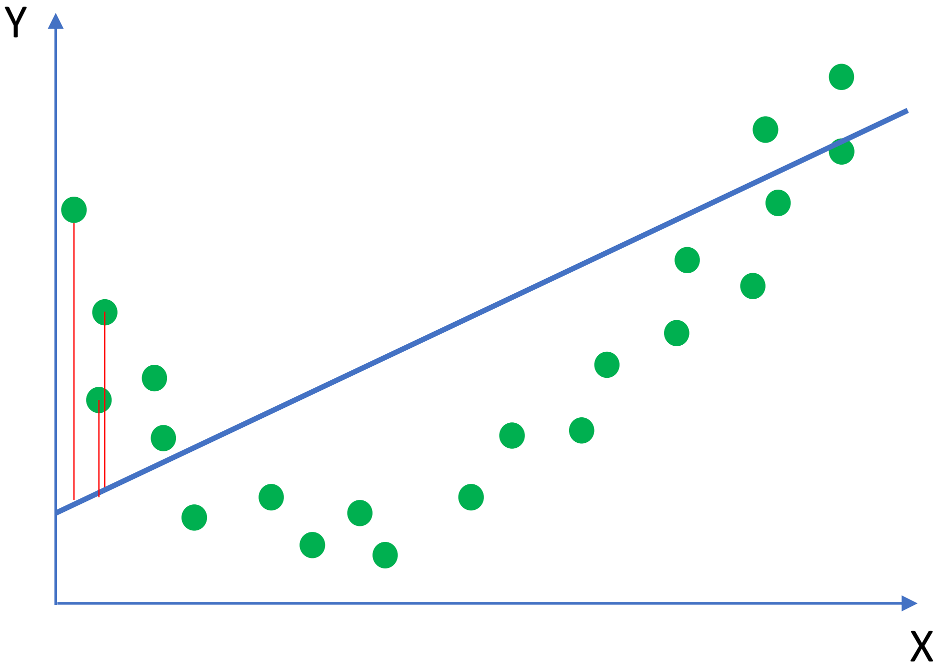
Bert Gollnick

Underfitting and Overfitting

Regression Example

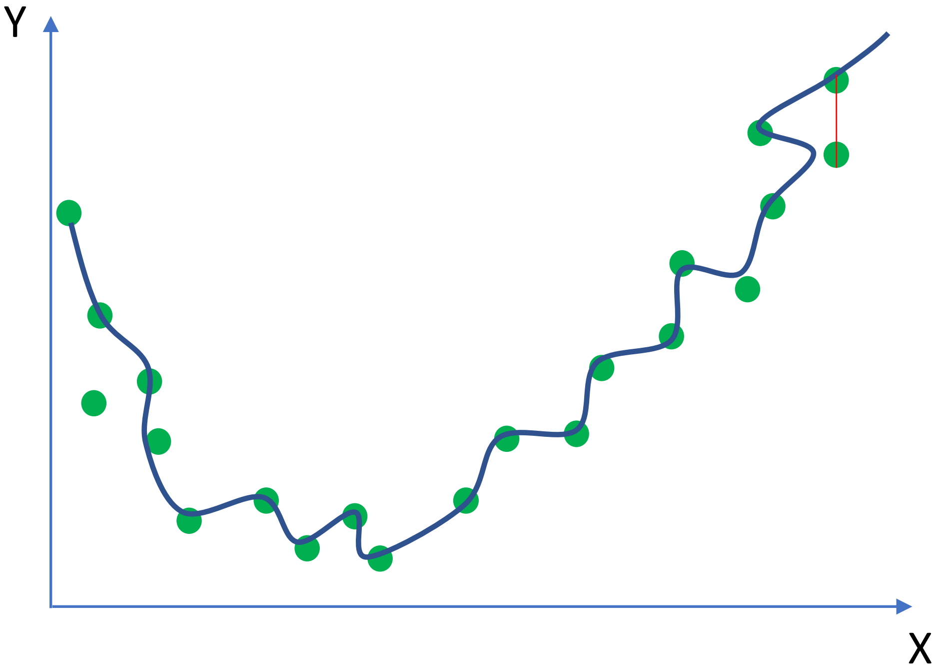
Underfitting

- High bias



Overfitting

- High variance

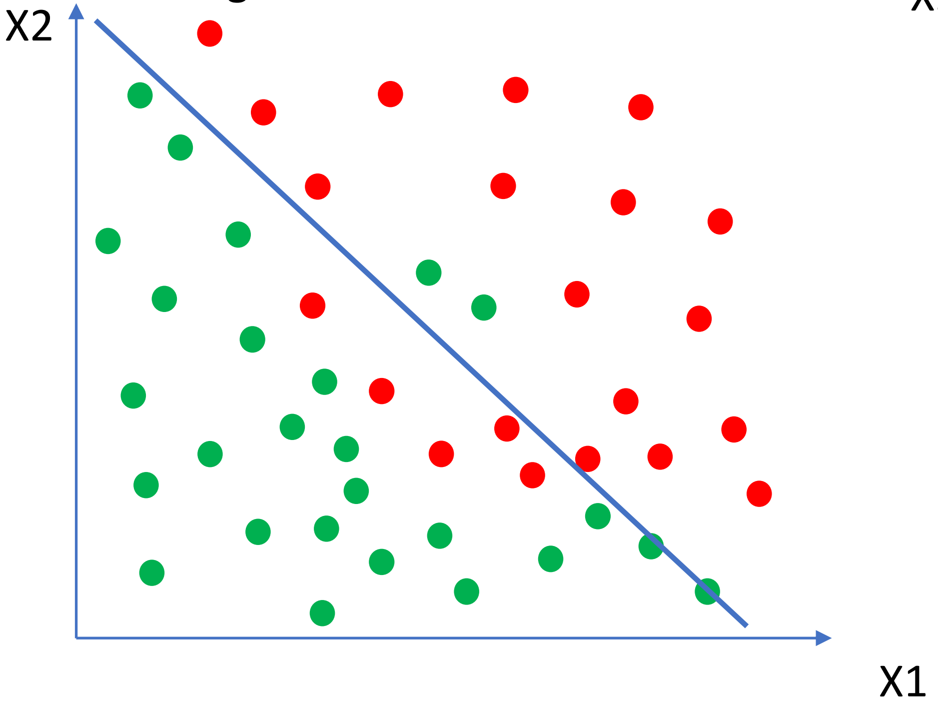


Underfitting and Overfitting

Classification Example

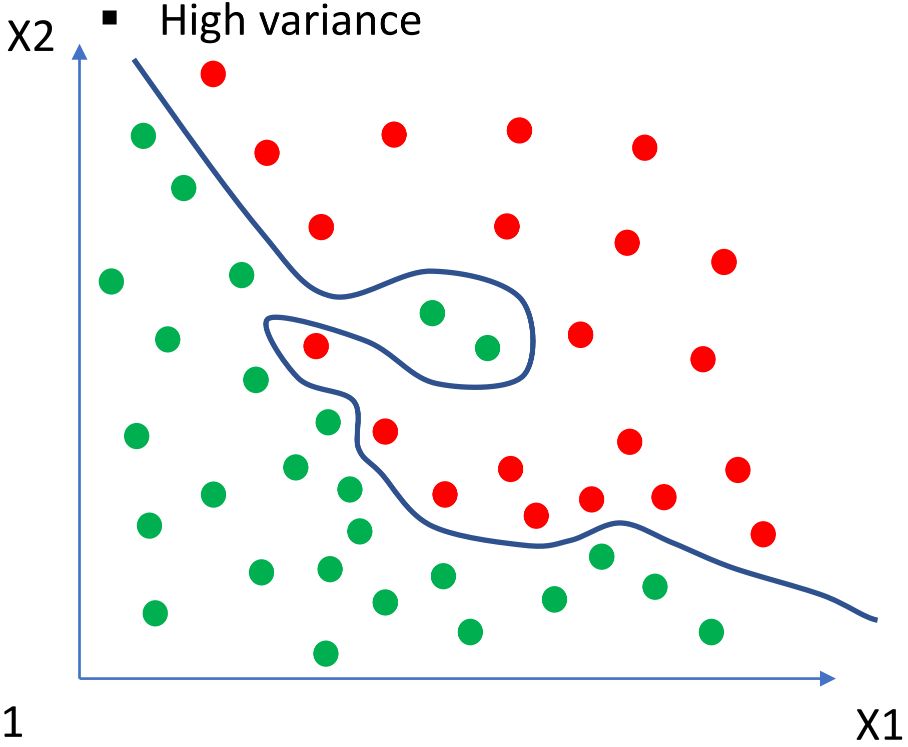
Underfitting

- High bias



Overfitting

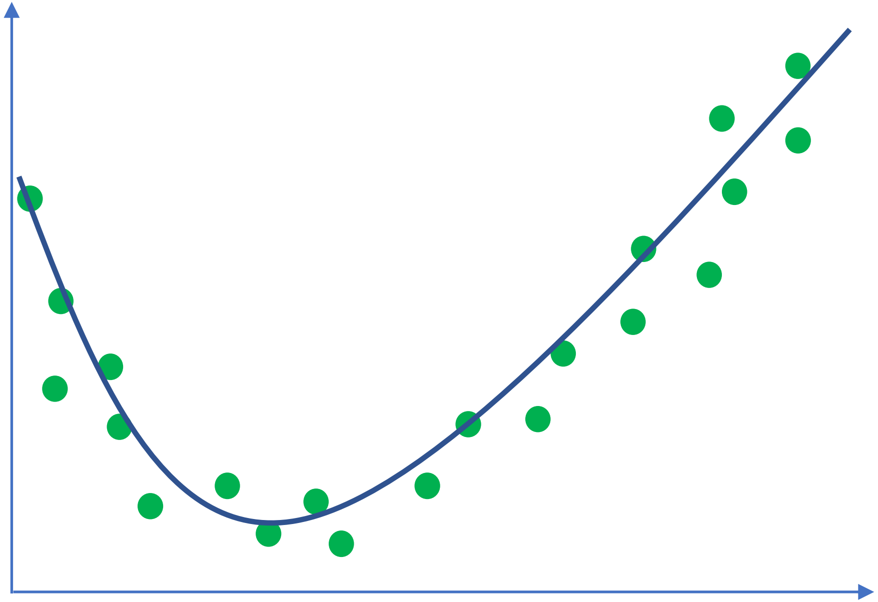
- High variance



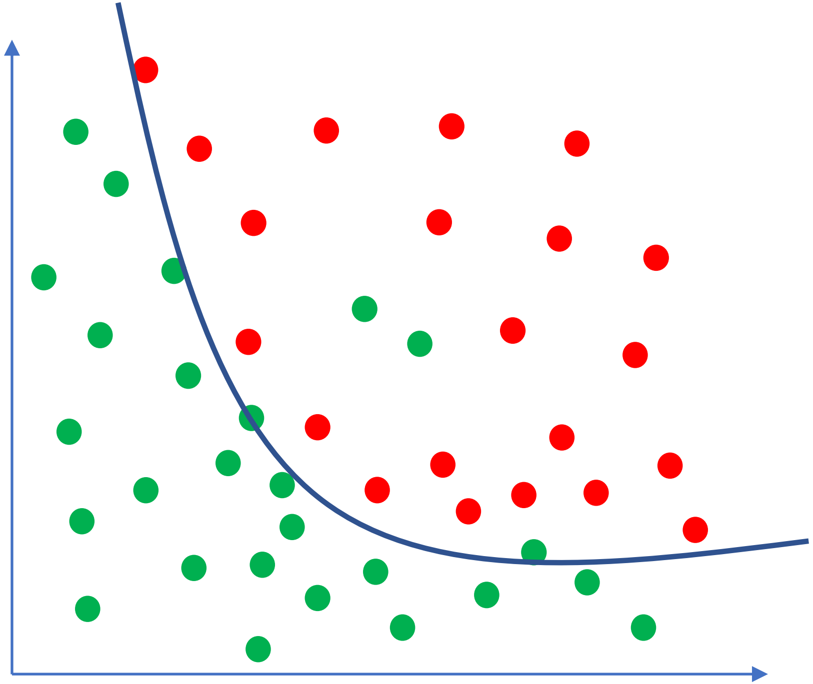
Underfitting and Overfitting

Good Fits

- Regression Example

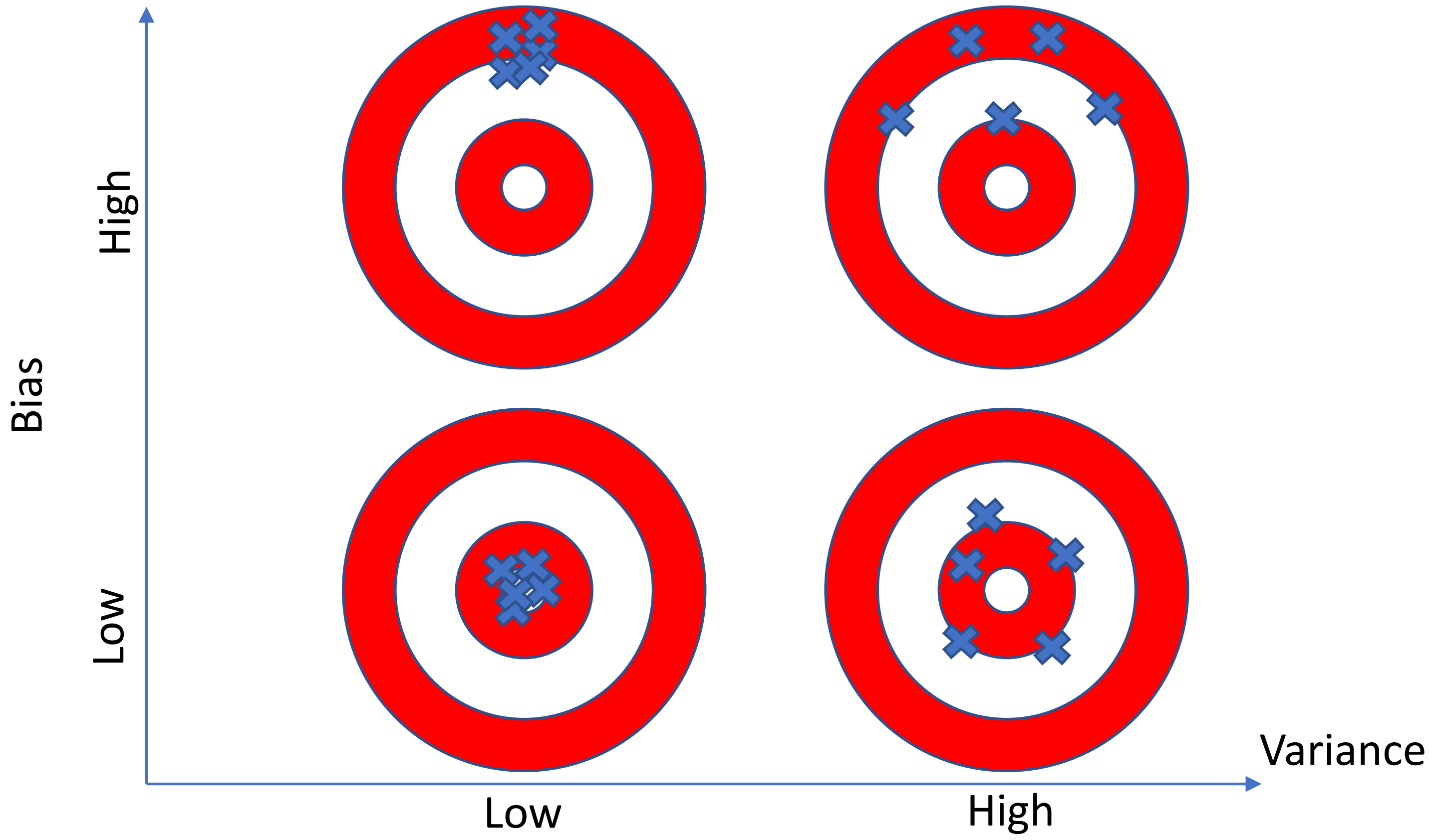


- Classification Example



Underfitting and Overfitting

Bias and Variance

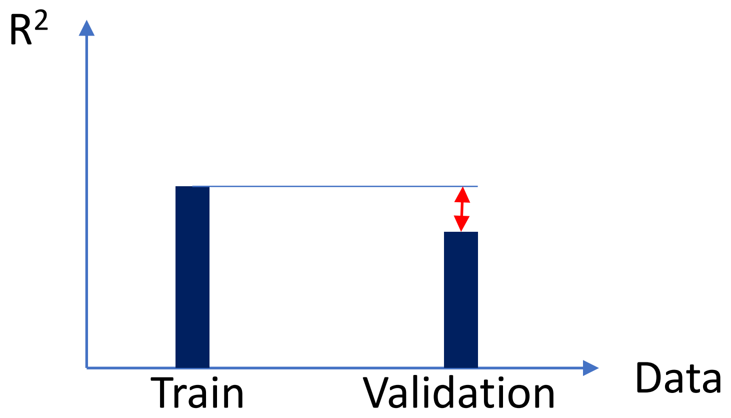


Underfitting and Overfitting

Bias and Variance

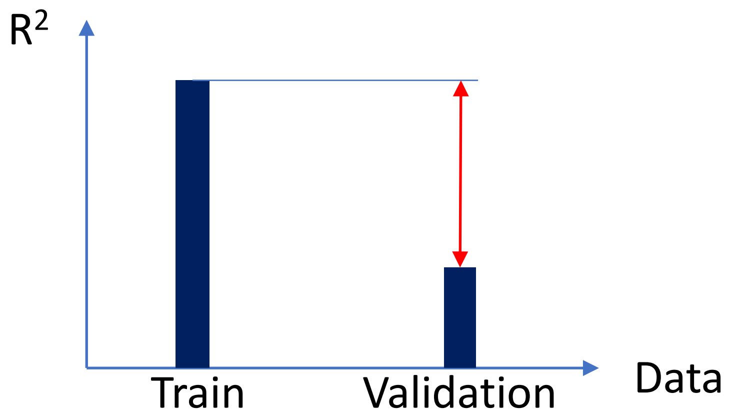
Bias

- Difference prediction / actual values of training data
- Example High Bias / Low Variance
  - R^2=0.5 for training
  - R^2=0.48 for validation



Variance

- Difference prediction of training data vs. prediction validation data
- Example Low Bias / High Variance
  - R^2=0.97 for training
  - R^2=0.3 for validation



Underfitting and Overfitting

Bias and Variance

Adding more parameters to a model increases model complexity

Model complexity ↑

Bias ↓

Variance ↑

Using a more complex model

Model complexity ↑

Bias ↓

Variance ↑

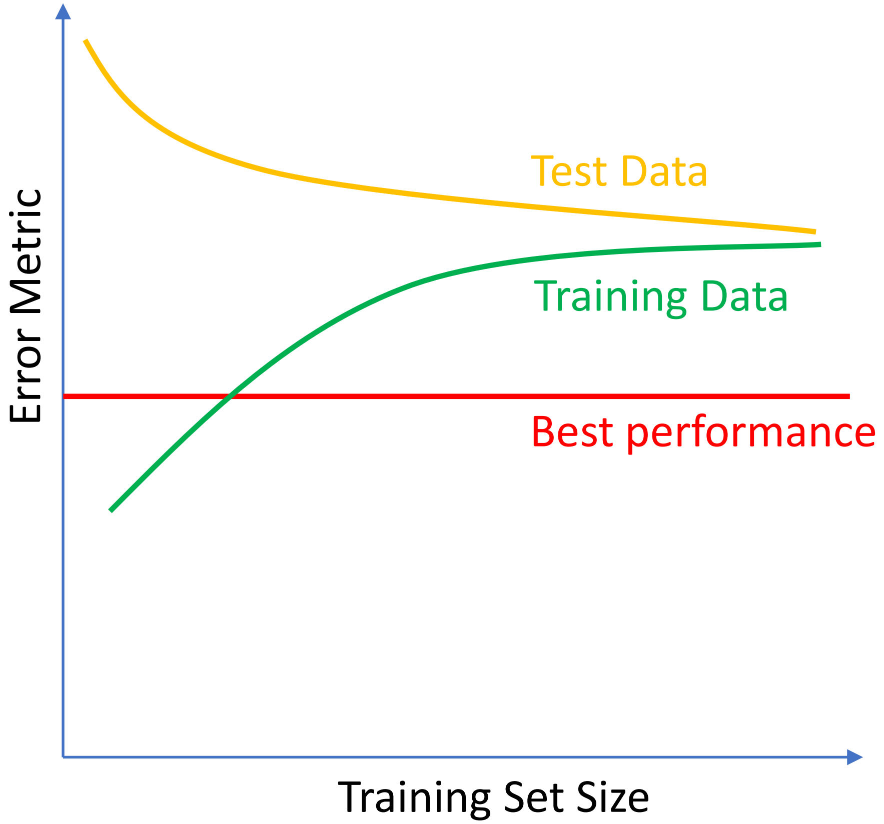
Underfitting and Overfitting

Bias

High Bias

- Learns fast
- Easy to understand (fewer parameters)
- Poor performance on complex problems

→ Underfitting

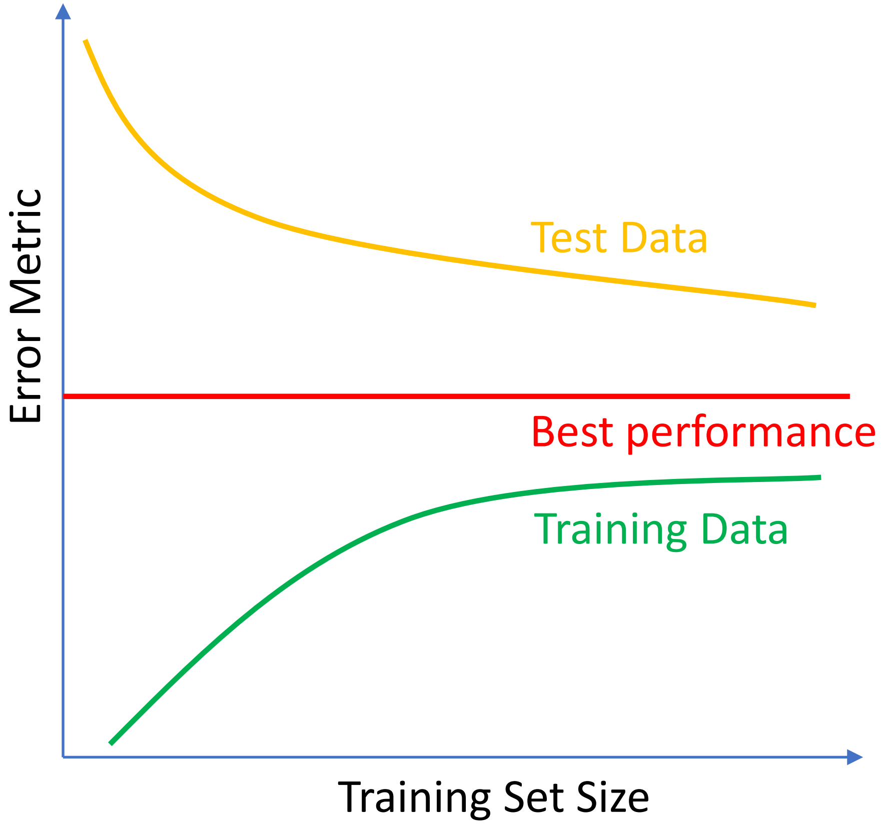


Underfitting and Overfitting

Variance

- High Variance → Large change of predictions if different training data is used
- Low variance algorithms: Linear Regression, LDA, Logistic Regression
- High variance algorithms: Decision Trees, kNN, SVM
- High Variance → more paramaters
- Good training performance, poor validation performance → poor generalization

→ Overfitting



Underfitting and Overfitting

Bias Variance Tradeoff

- Goal:
  - low bias / low variance
  - Good prediction performance
- Bias and variance have opposite directions
- Linear ML algorithms typically high bias, low variance
- Non-linear ML algorithms typically low bias, high variance
- Non-linear ML algorithms often have hyperparameters for tuning