

Bias-Variance Trade-Off

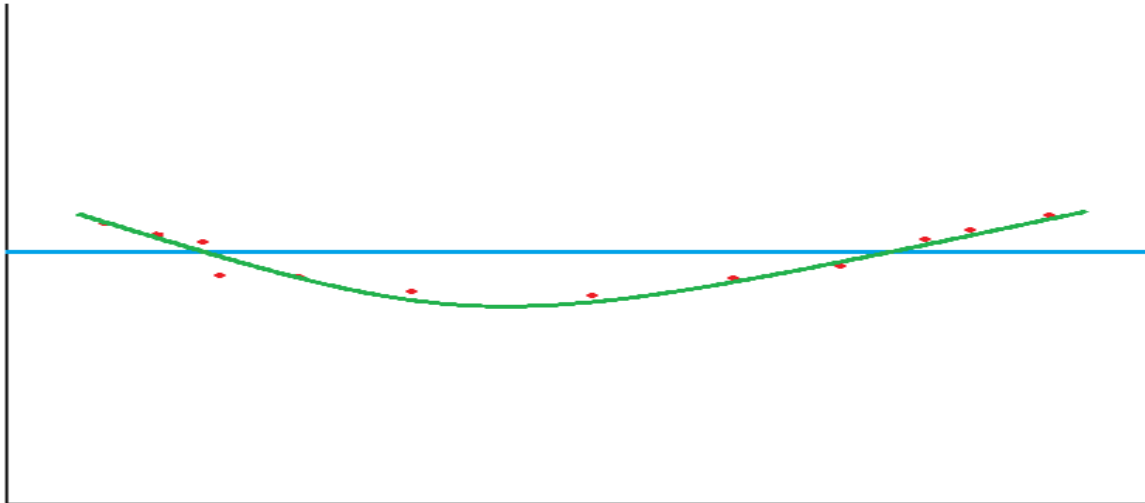
As a Concept

Error Causes

- There are two causes of error which is reducible
 - Bias
 - Variance

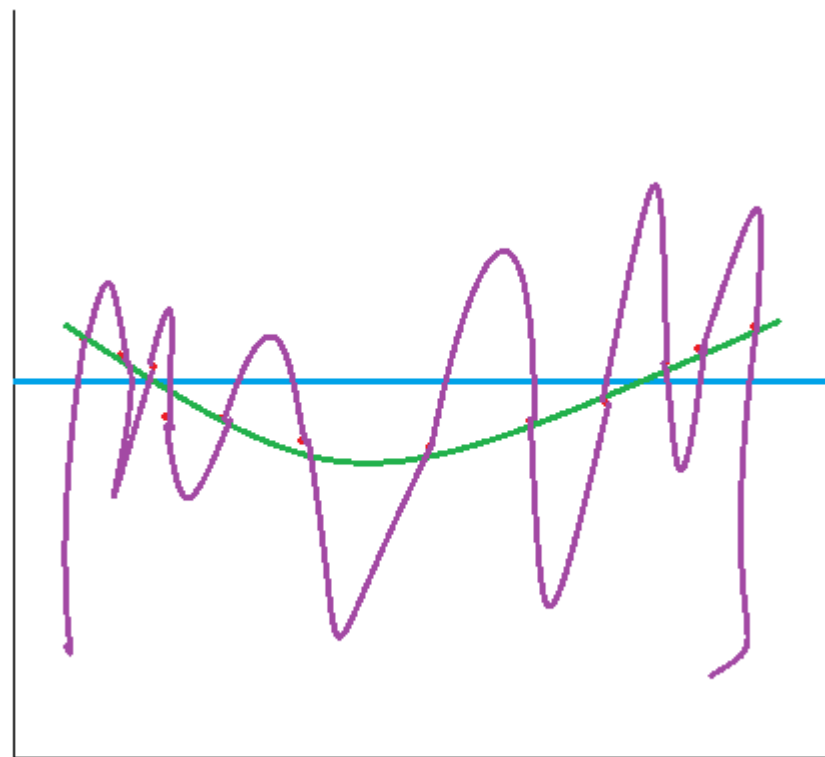
Bias

- Error due to bias is the error in prediction due to wrong assumptions about the data
- Due to wrong assumptions a wrong model may be applied e.g. if data is non-linear and assuming as linear we apply linear regression
- More are the restrictions on the model, more is the bias higher



Variance

- Error due to variance is error due to sampling of the training set
- Model having high variance fits closely to the training set
- Model in this case has less restrictions
- If any of the training set observation changes, then the model changes completely
- Higher variance model predicts on the test badly



Bias Variance Trade -Off

- Lower is the Bias, Higher is the Variance
- Lower is the Variance, Higher is the Bias

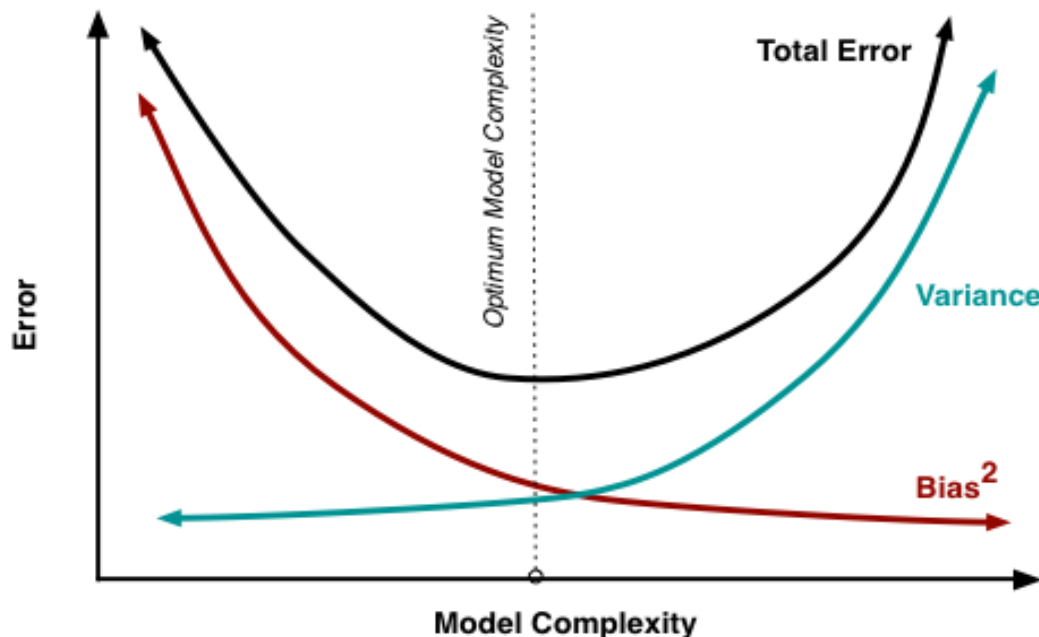


Image Courtesy: http://www.bogotobogo.com/python/scikit-learn/scikit_machine_learning_Bias-variance-Tradeoff.php

Bias-Variance Trade-Off

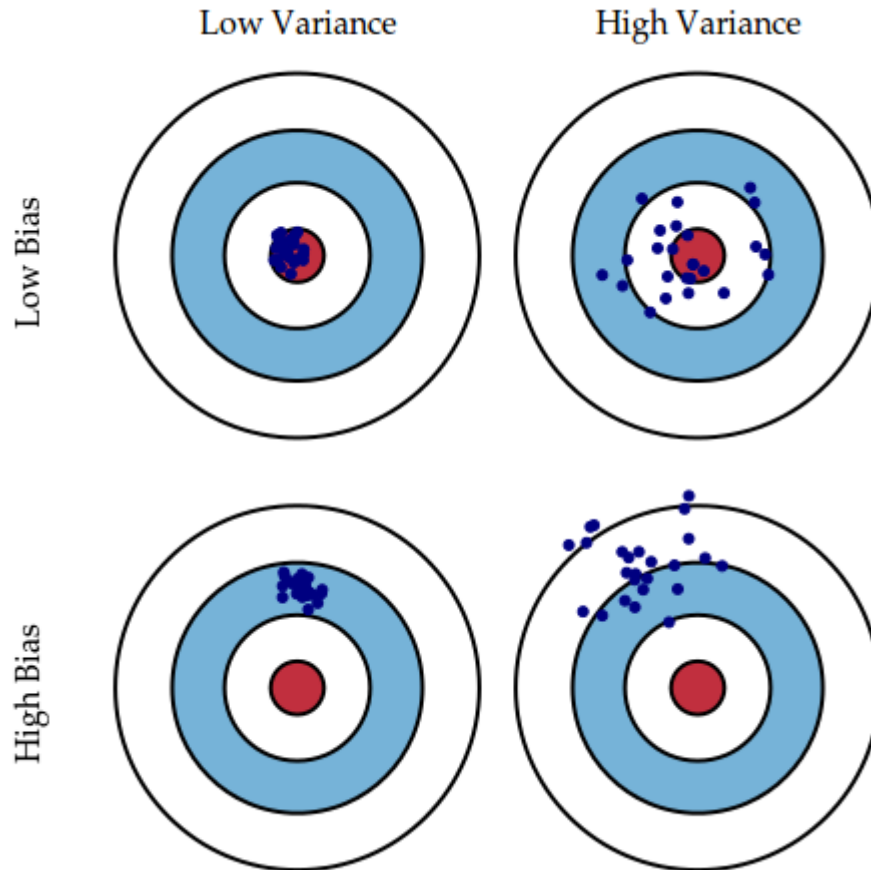


Image Courtesy: <http://scott.fortmann-roe.com/docs/BiasVariance.html>

Overfitting and Underfitting

Overfitting

- Model fits the training data very well but its too specific to the training set
- Has High Variance
- Results in poor accuracy ultimately

Underfitting

- Model is restricted and hence doesn't fit properly to the training set
- Has High Bias
- Results in poor accuracy ultimately