Bias vs. Variance

Week 04 - Day 04

Interesting Concept

Complete different approach!

Overfitting vs. Underfitting

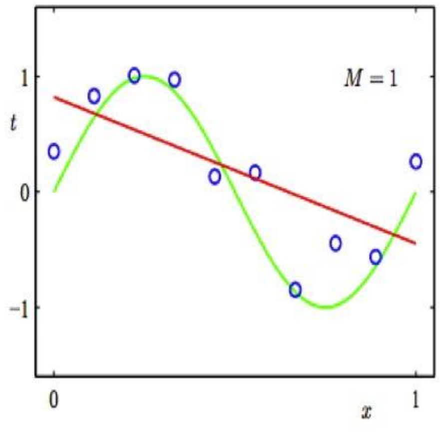
Underfitting = ???

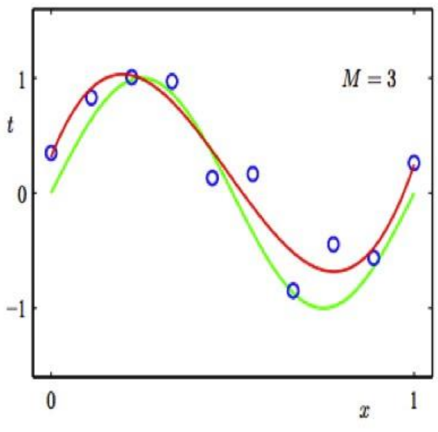
Overfitting = ???

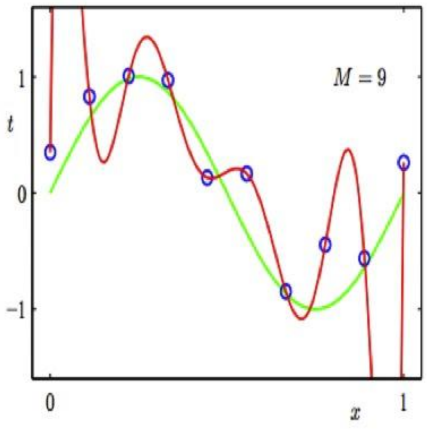
Underfitting = model is too simple

Overfitting = model is too complex

Under- and Over-fitting examples







predictor too inflexible:
cannot capture pattern

predictor too flexible:
fits noise in the data

Overfitting

1. Model is good at describing the data
2. Model is terrible at predicting new unseen data

Overfitting

1. Model is good at describing the data
2. Model is terrible at predicting new unseen data

generalising

We want: model the signal

We don't want: learn the noise

Linear Regression
Vs.
Polynom. Regression

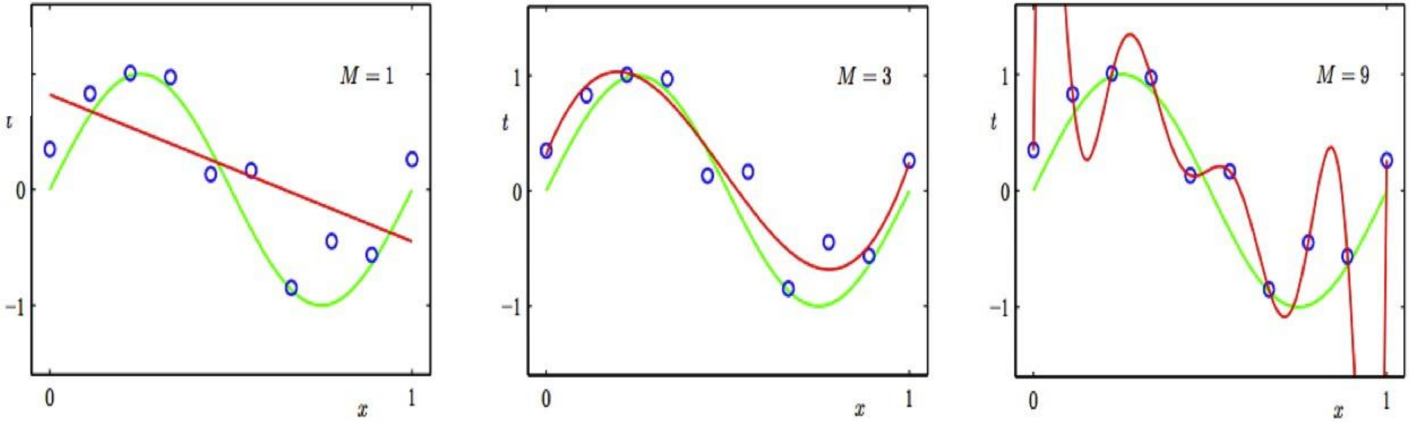
https://arachnoid.com/polysolve/

Different Samples

(e.g. polls in different countries)

++ Error
-- Variance

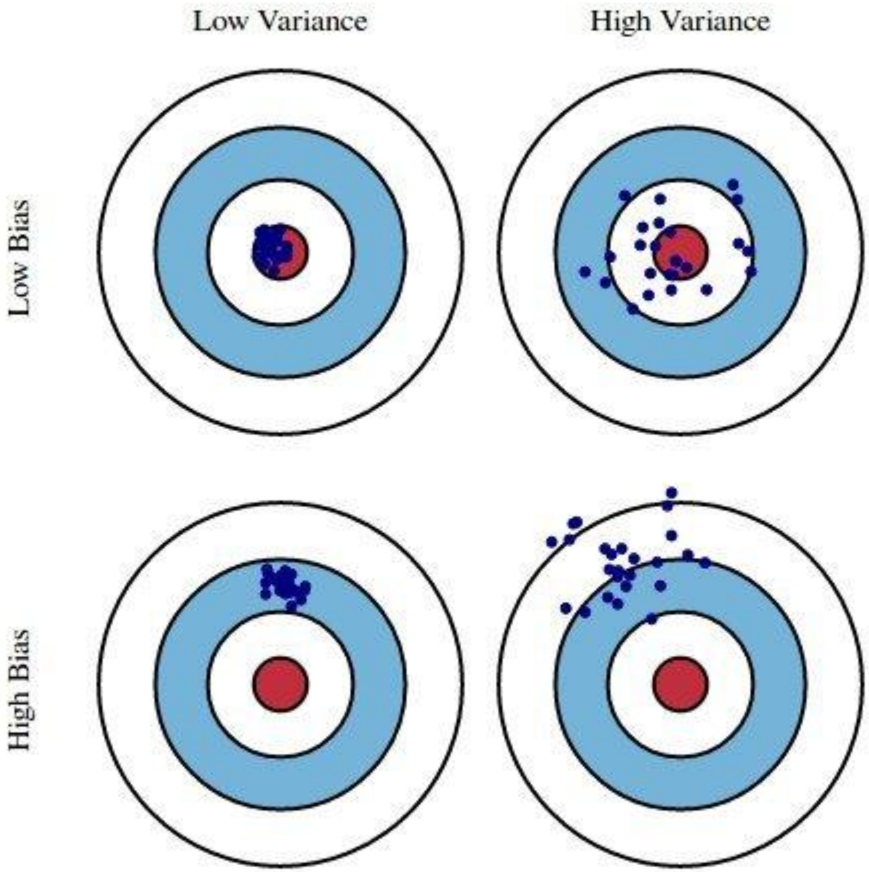
-- Error
++ Variance

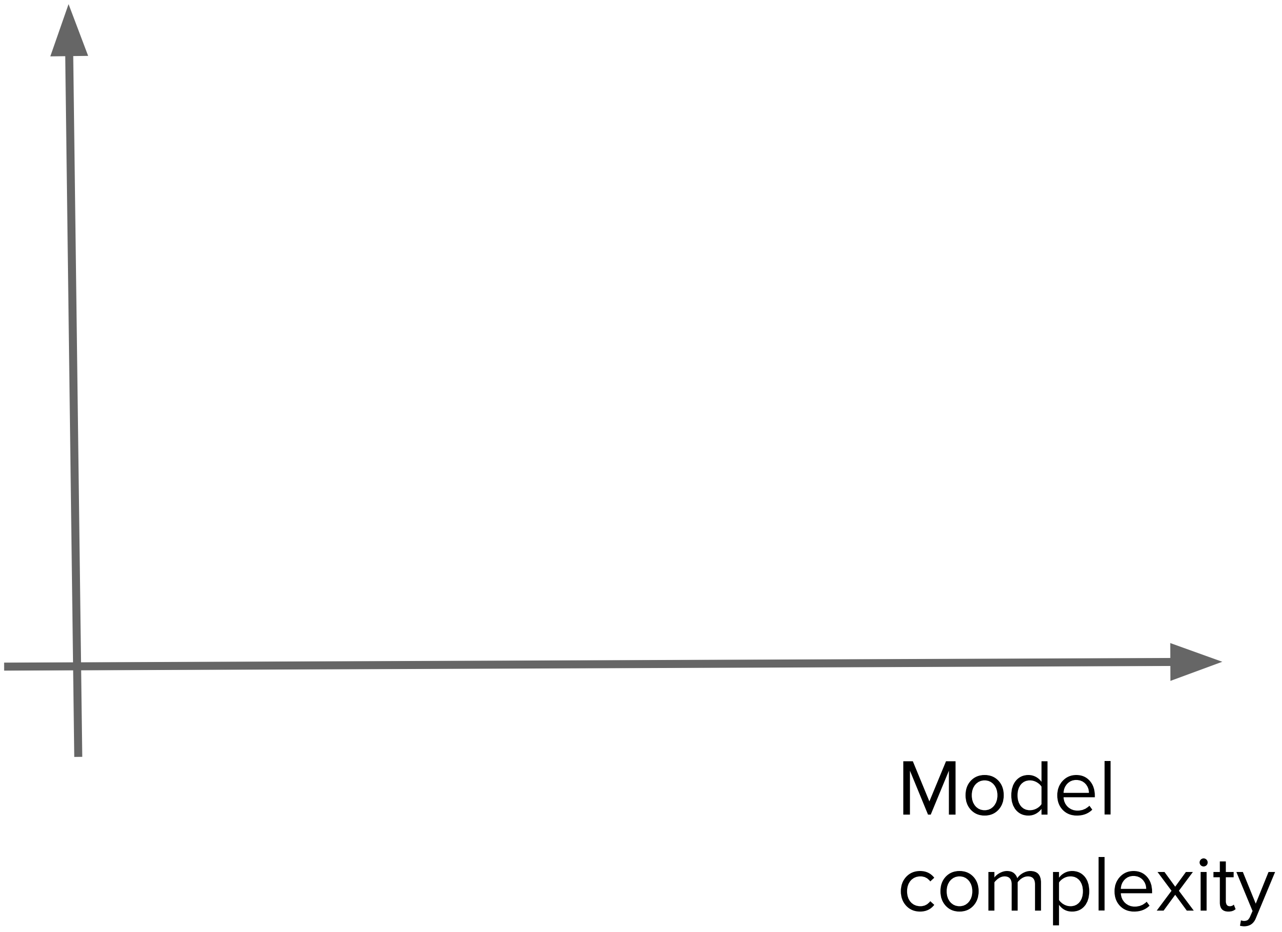


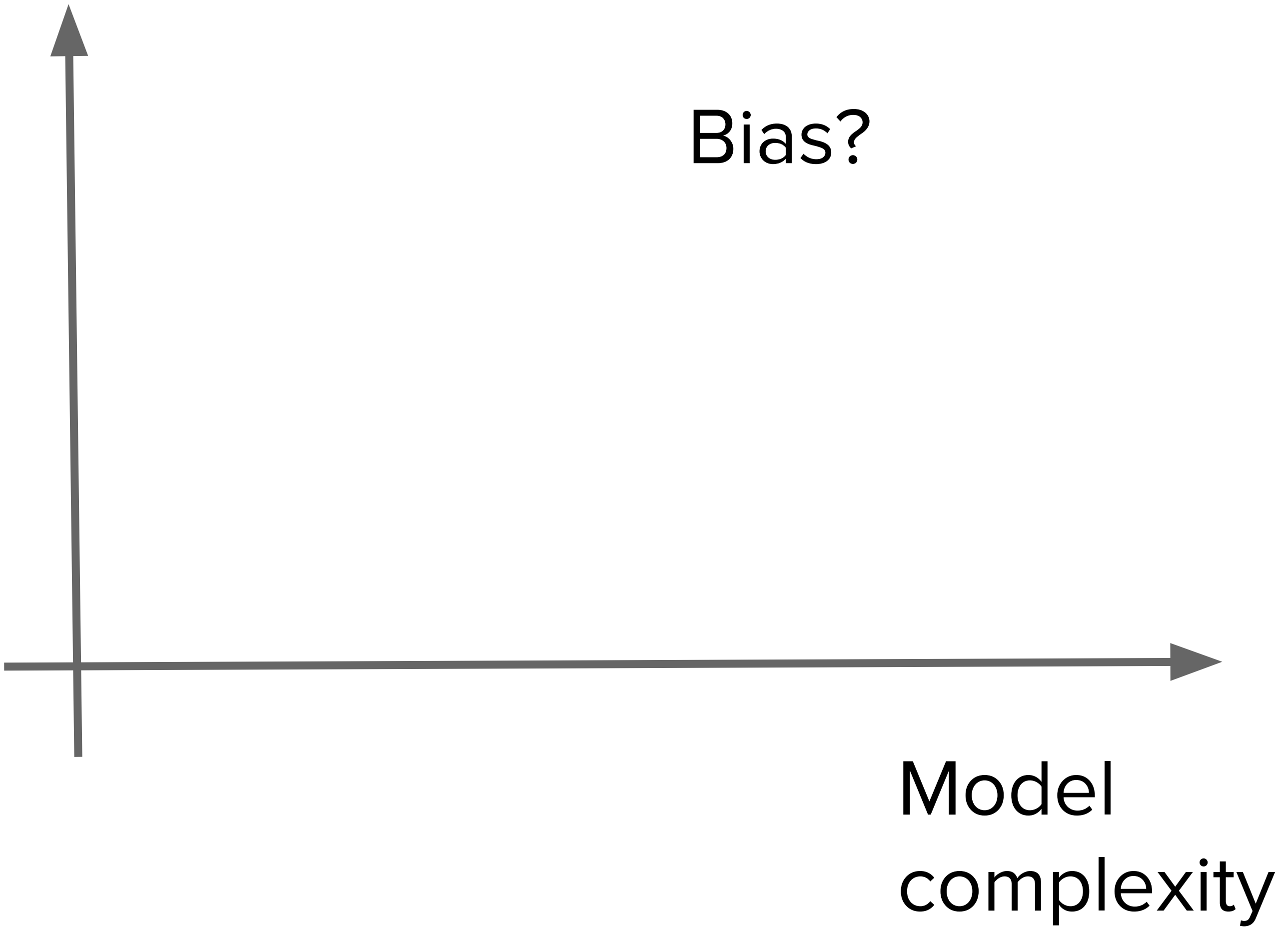
Bias vs. Variance

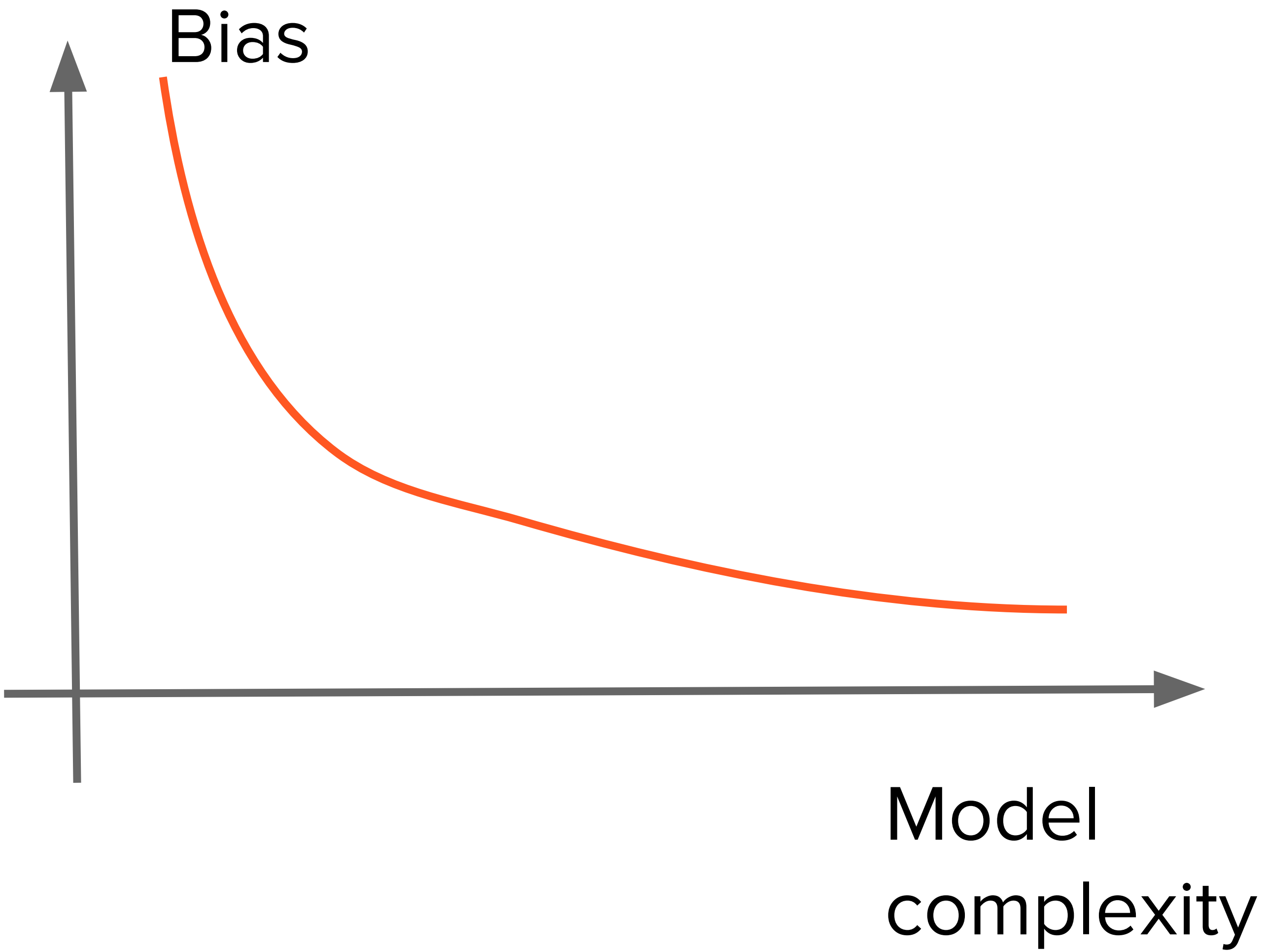
Bias = error because the model is too easy

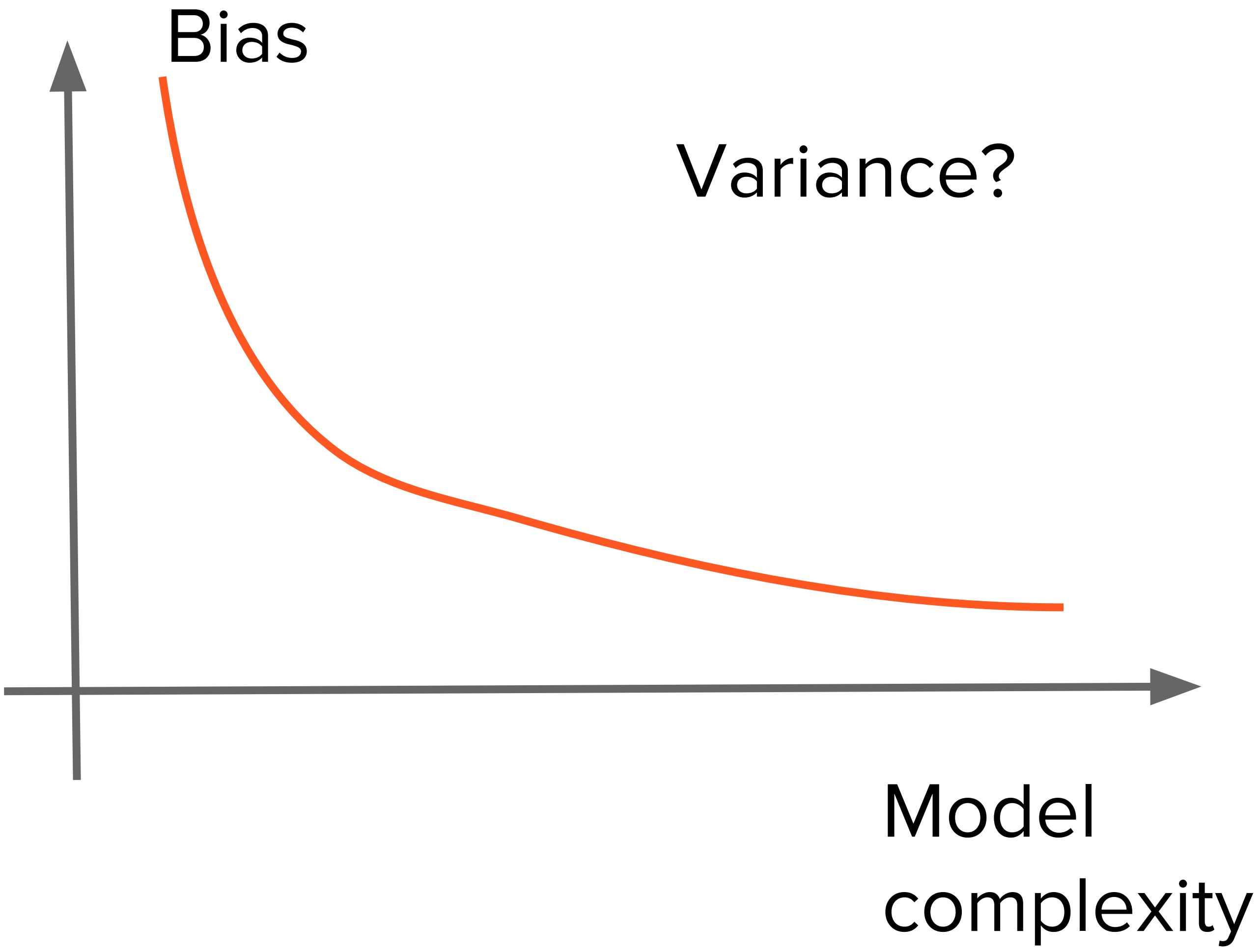
Variance = error because the model is too complex

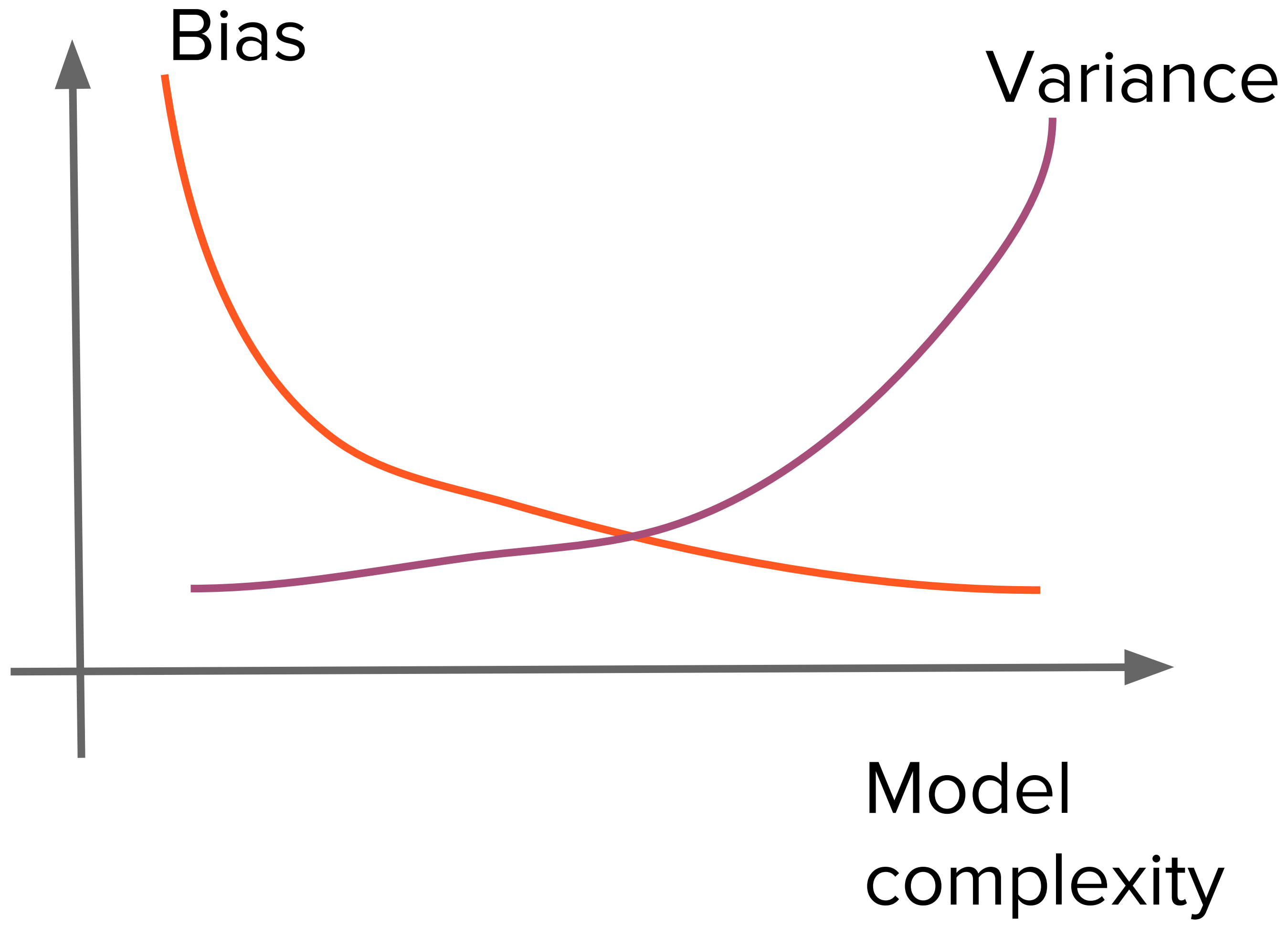

Low Variance
High Variance
Low Bias
High Bias

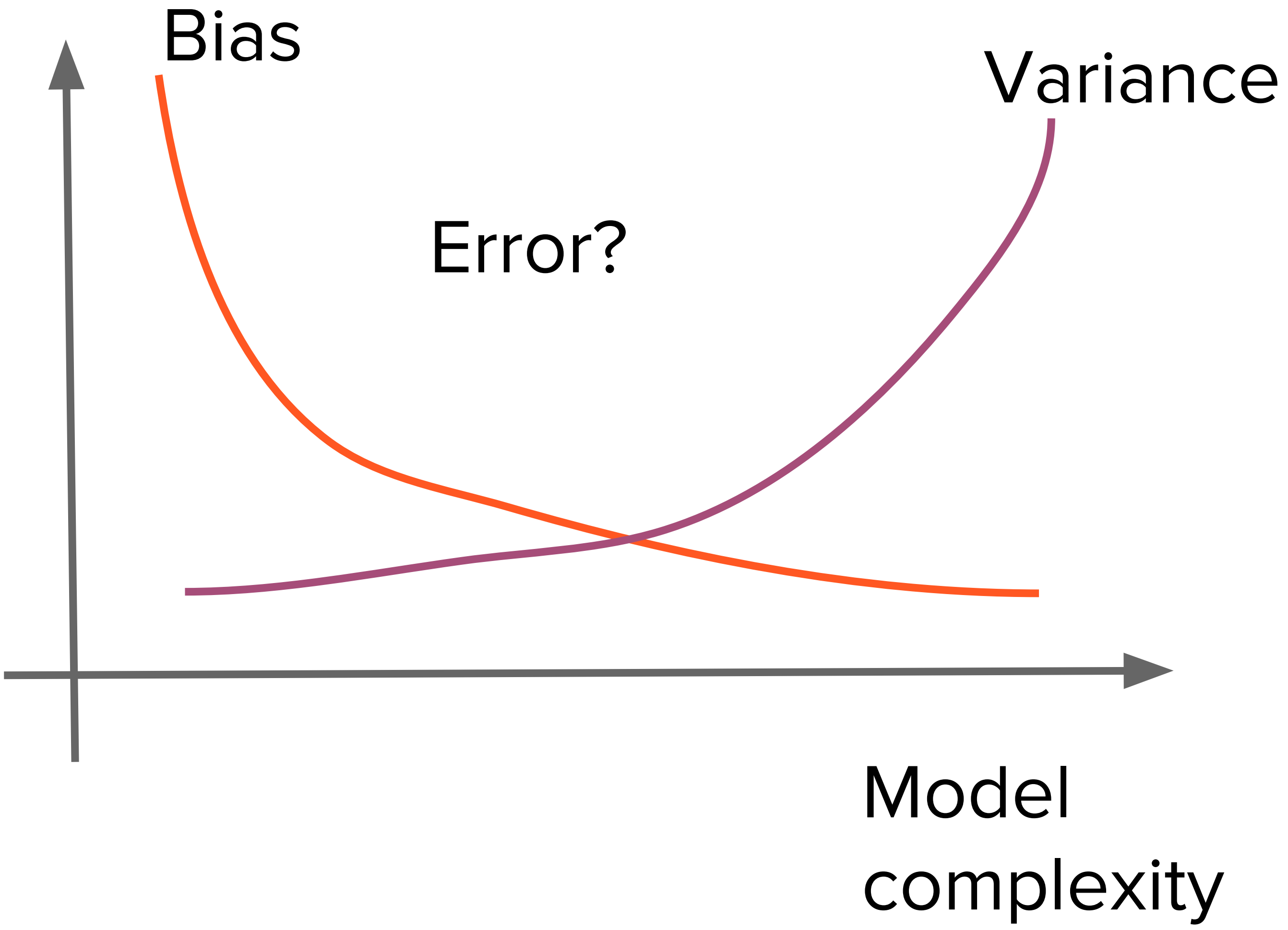

Model
complexity

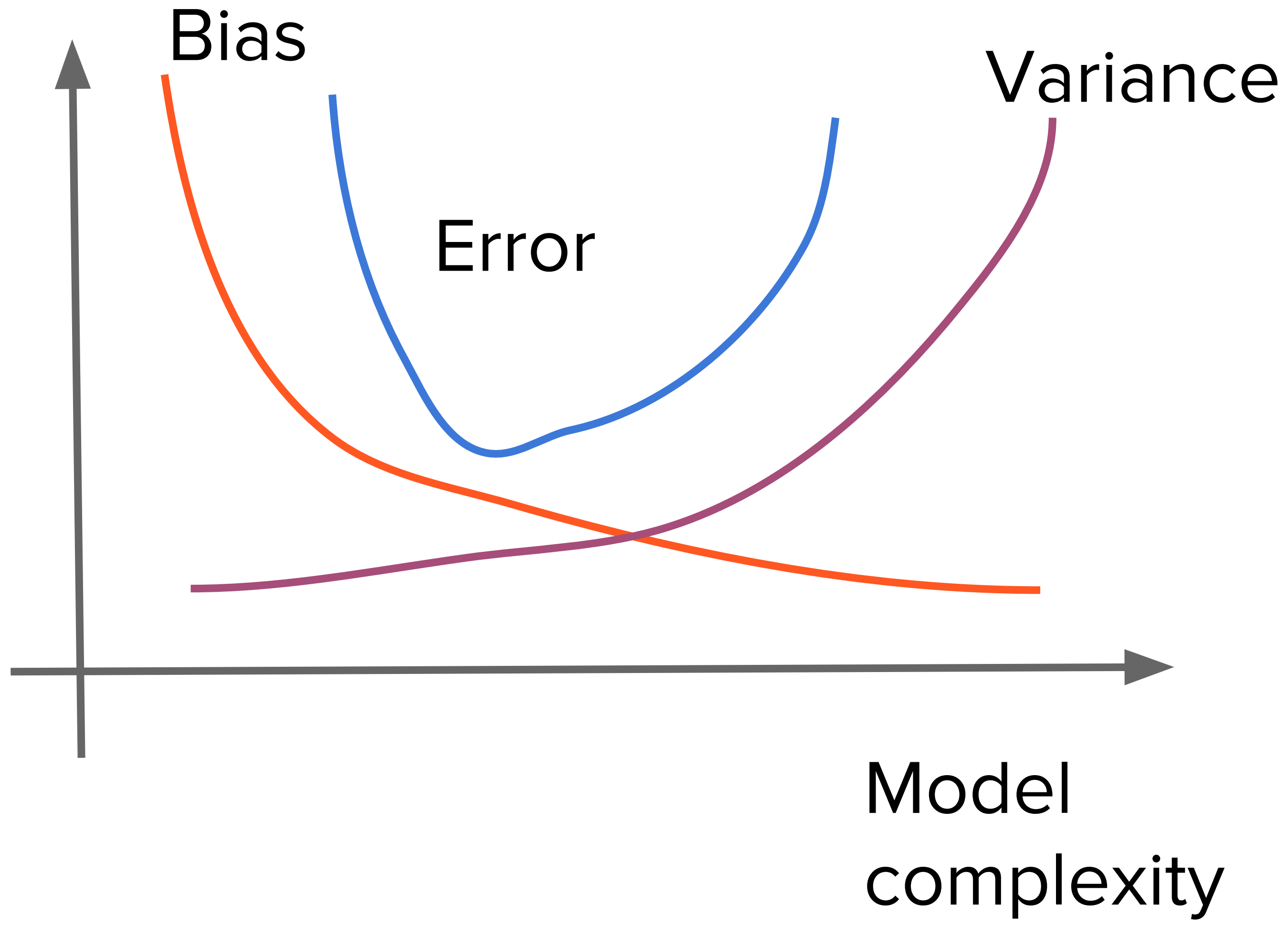

Bias?
Model
complexity

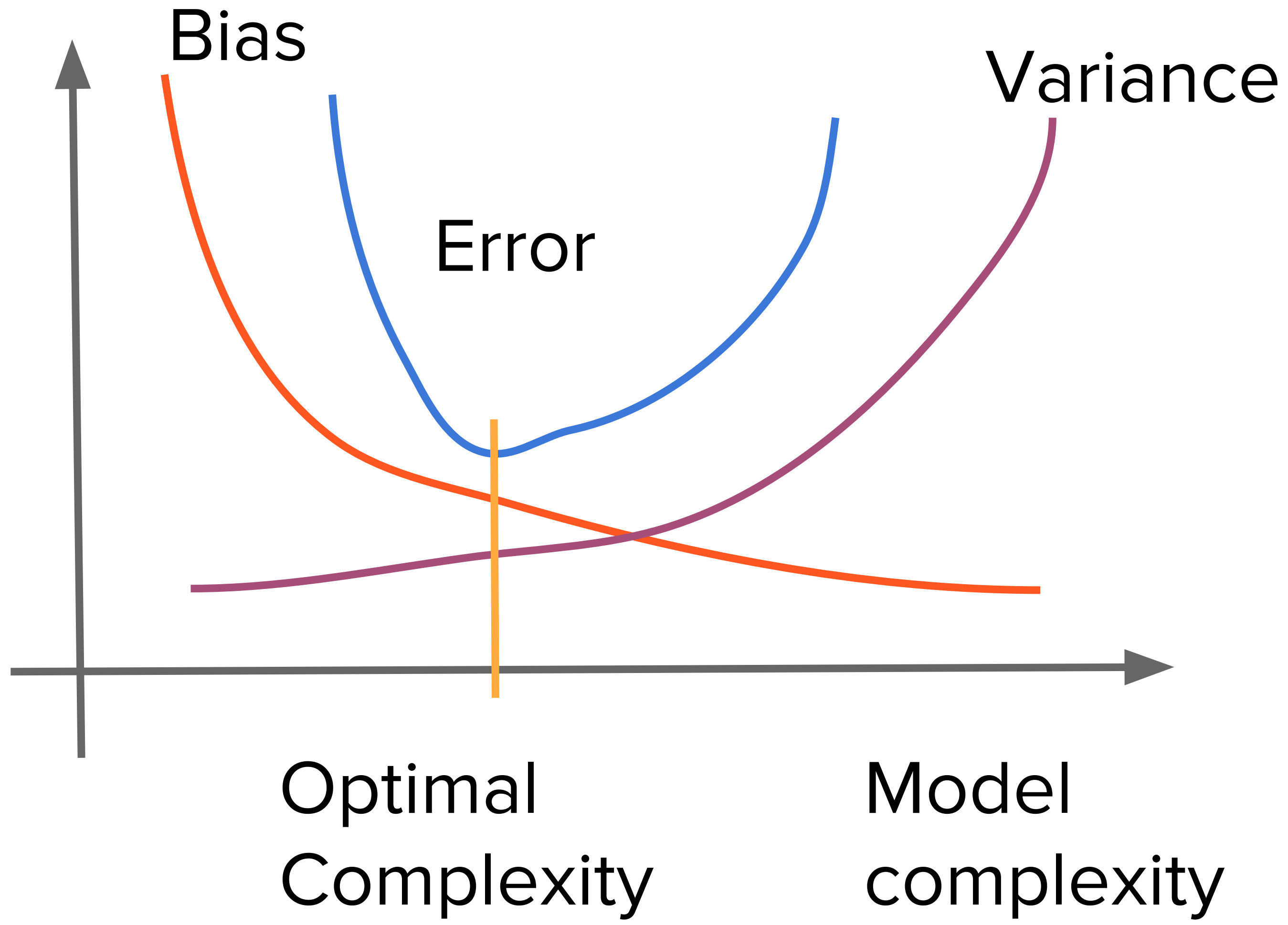

Bias
Model
complexity


Bias
Variance?
Model
complexity


Bias
Variance
Model
complexity


Bias
Variance
Error?
Model
complexity


Bias
Variance
Error
Model
complexity


Bias
Variance
Error
Optimal
Complexity
Model
complexity

Mathematical Decomposition

$$\text{Error} = \text{Bias}^2 + \text{Variance} + \text{Irreducible Error}$$

$$E\left[\left(\hat{f}(x) - f(x)\right)^2\right] = \left(\mathrm{E}[\hat{f}(x)] - f(x)\right)^2 + \mathrm{E}\left[(\hat{f}(x) - \mathrm{E}[\hat{f}(x)])^2\right] + E\left[\left(y - f(x)\right)^2\right]$$

Practical Meaning?

Interesting Concept

Complete different approach!

Interesting Concept

~~Complete different approach!~~

Trivial practical meaning

1) Always use cross validation!
2) Try different models