Bias-Variance Trade-off

Natasha Balac, Ph.D.

Bias and Variance

- No “best classifier” in general
  - Necessity for exploring a variety of methods
- How to evaluate if the learning algorithm “matches” the classification problem
- Bias: measures the quality of the match
  - High-bias implies *poor* match
  - Error rate of a particular learning algorithm
- Variance: measures the specificity of the match
  - High-variance implies a *weak* match
  - Comes from a particular data set used
- Bias and variance are not independent of each other

Bias Variance Dilemma

- Procedures with increased flexibility to adapt to training data have lower bias, but higher variance
  - Large number of parameters
  - Fits well and have low bias, but high variance
- Inflexible procedures have higher bias, but lower variance
  - Fewer number of parameters
  - May not fit well to data: have high bias, but low variance
- A large amount of training data generally helps improve performance of estimation if the model is sufficiently general to represent the target function

Model Loss (Error)

- Squared loss of model on test case i:

$$\left(Learn(x_i, D) - Truth(x_i)\right)^2$$

- Expected prediction error:

$$\left\langle \left(Learn(x, D) - Truth(x)\right)^2 \right\rangle_D$$

Bias/Variance Decomposition

$$\left\langle (L(x,D) - T(x))^2 \right\rangle_D = Noise^2 + Bias^2 + Variance$$

$Noise^2$ = lower bound on performance

$Bias^2$ = (expected error due to model mismatch)^2

Variance = variation due to train sample and randomization

Bias

- Low bias
  - linear regression applied to linear data
  - 2nd degree polynomial applied to quadratic data
  - ANN with many hidden units trained to completion
- High bias
  - constant function
  - linear regression applied to non-linear data
  - ANN with few hidden units applied to non-linear data

Variance

- Low variance
  - constant function
  - model independent of training data
  - model depends on stable measures of data
    - mean
    - median
- High variance
  - high degree polynomial
  - ANN with many hidden units trained to completion

Sources of Variance in Supervised Learning

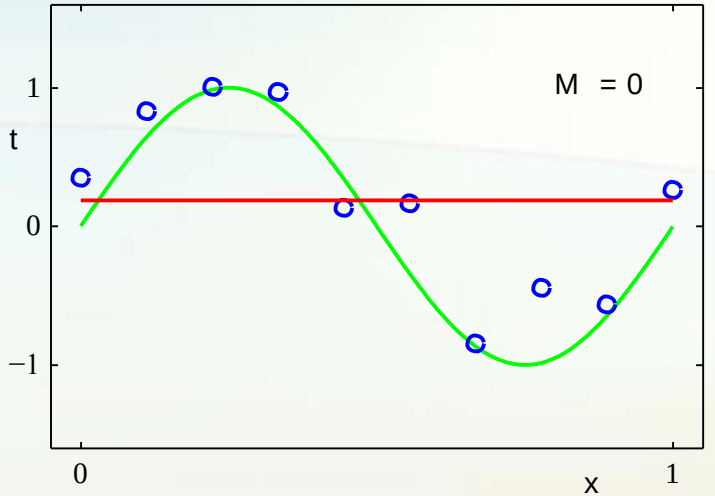
- noise in targets or input attributes
- bias (model mismatch)
- training sample
- randomness in learning algorithm
  - neural net weight initialization
- randomized subsetting of train set:
  - cross validation, train and early stopping set

Bias/Variance Tradeoff

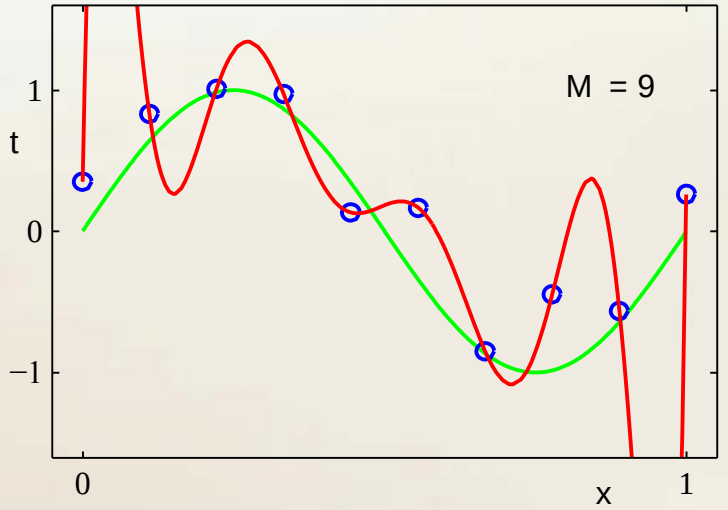
- (bias^2+variance) is what counts for prediction
- Often:
  - low bias => high variance
  - low variance => high bias
- Tradeoff:
  - bias^2 vs. variance

Bias-Variance Trade-off

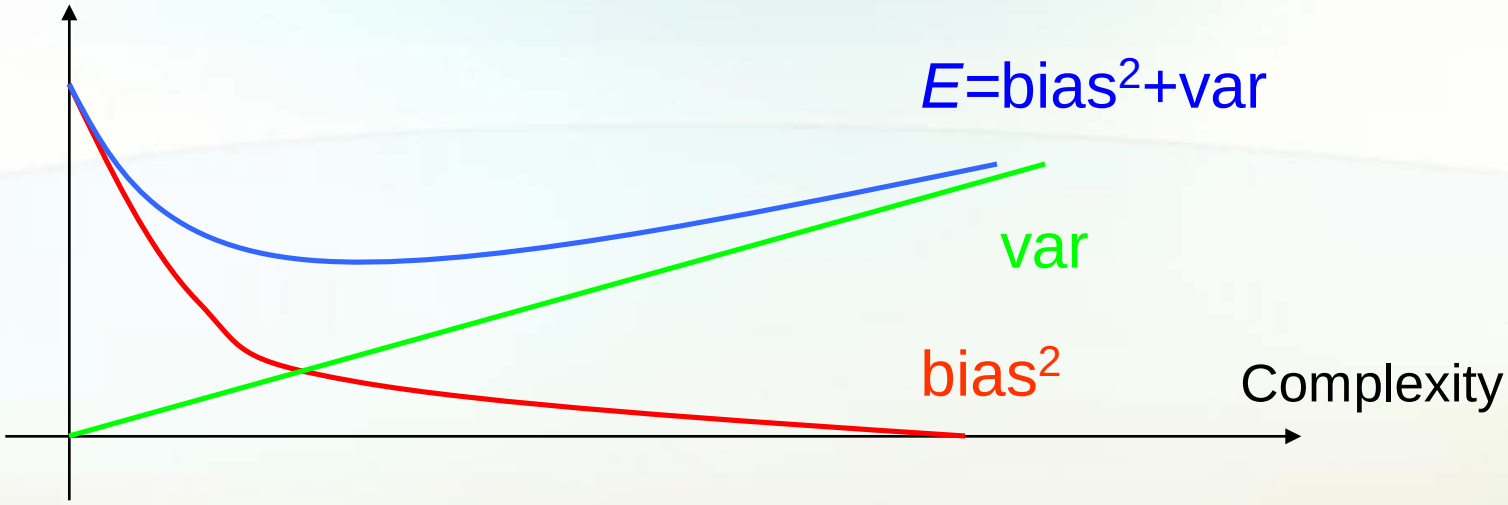
- Model too simple: does not fit the data well
  - A *biased* solution



- Model too complex: small changes to the data, solution changes a lot
  - A *high-variance* soluJon



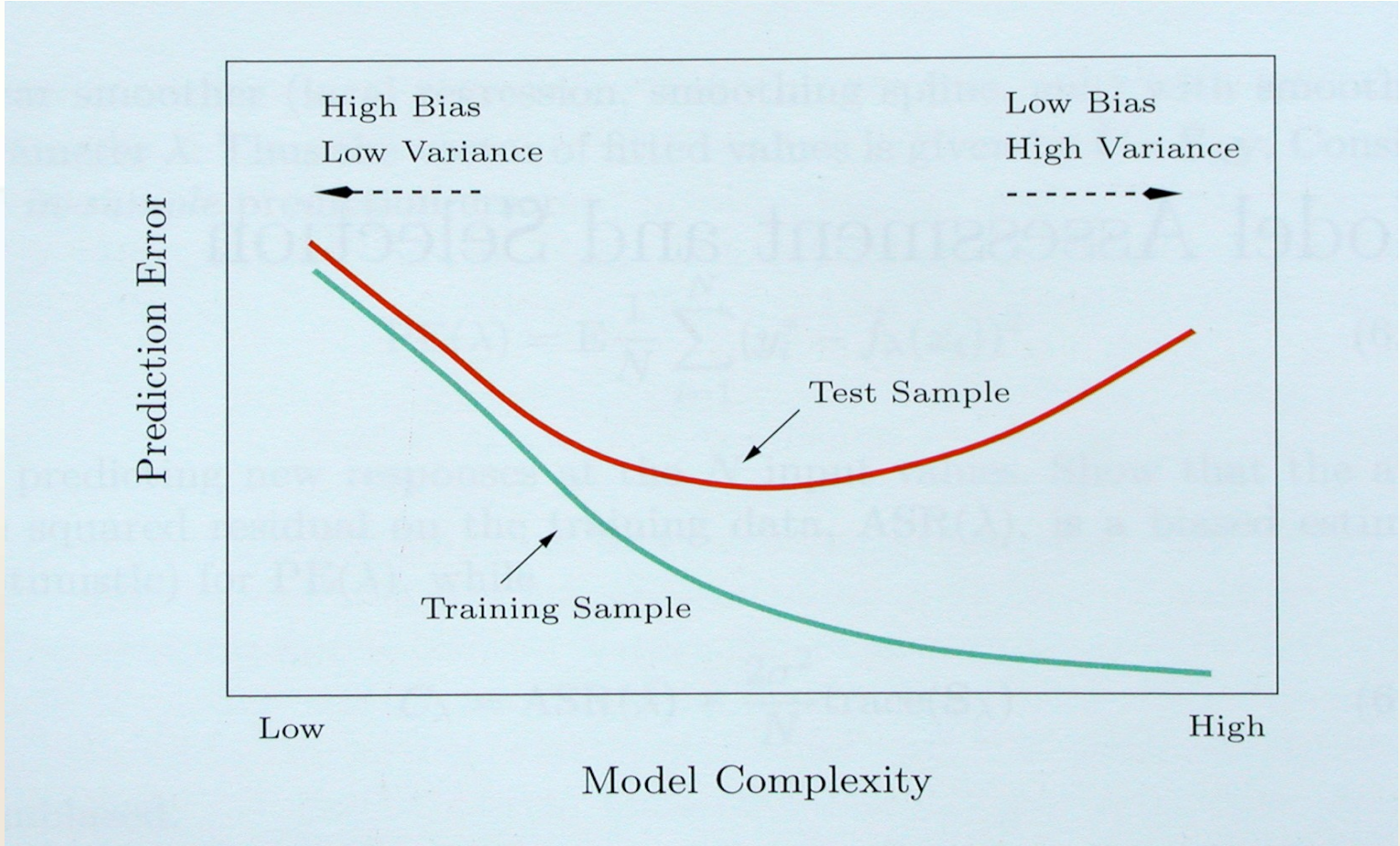
Complexity of the model



Typically the bias is a decreasing function of the complexity

while variance is an increasing function of the complexity

Bias/Variance Tradeoff



Hastie, Tibshirani, Friedman "Elements of Statistical Learning"
2001

Reduce Variance Without Increasing Bias

- Averaging reduces variance:

$$Var(\overline{X}) = \frac{Var(X)}{N}$$

- Average models to reduce model variance
- One problem:
  - only one train set
  - where do multiple models come from?

Bias Variance Dilemma

- Bias/Variance considerations recommend that we gather as much prior information about the problem as possible to find a best match for the classifier, and as large a dataset as possible to reduce the variance
- We can virtually never get zero bias and zero variance

Questions?

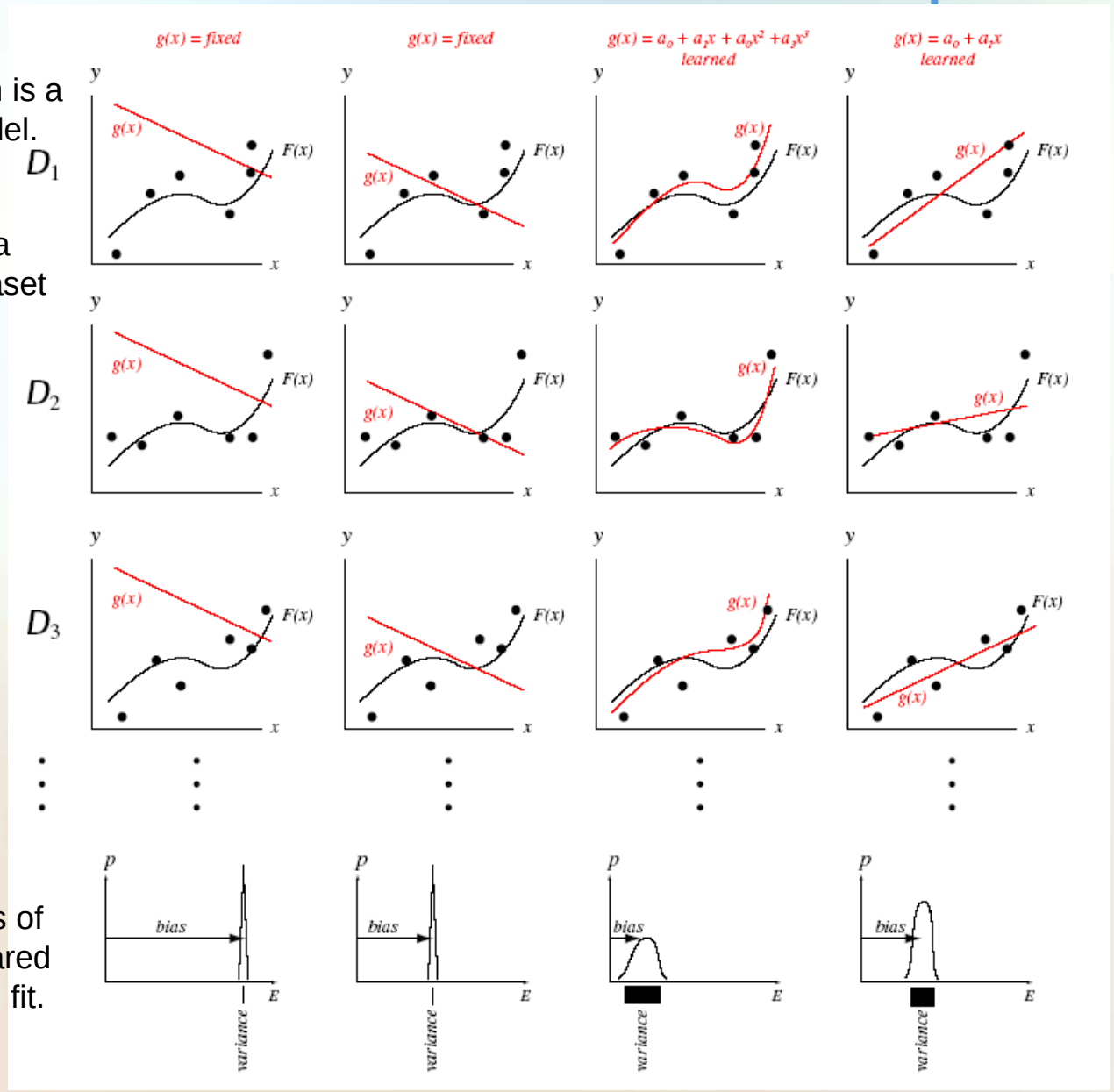
Thank you!

Bias-Variance Dilemma Example

Each column is a different model.

Each row is a different dataset of 6 points.

Histograms of mean-squared error of the fit.



Col 1:
Poor fixed linear model;
High bias, zero variance

Col 2:
Slightly better fixed linear model;
Lower (but high) bias, zero variance.

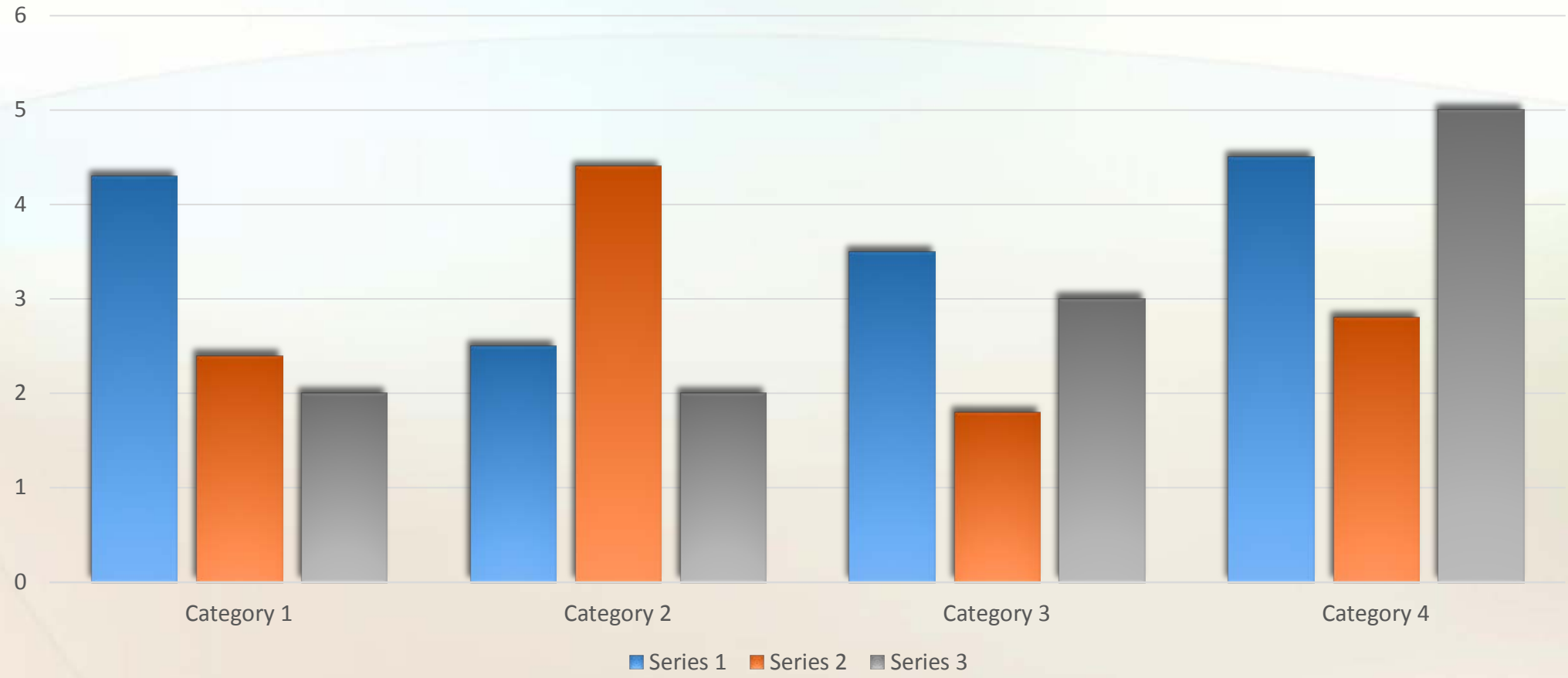
Col 3:
Learned cubic model;
Low bias, moderate variance.

Col 4:
Learned linear model;
Intermediate bias and variance.

Title and Content Layout with List

- Click to edit Master text styles
  - Second level
    - Third level
      - Fourth level
        - Fifth level

Title and Content Layout with Chart



Two Content Layout with Table

	Group A	Group B
Class 1	82	85
Class 2	76	88
Class 3	84	90

- First bullet point here
- Second bullet point here
- Third bullet point here

Two Content Layout with SmartArt

- First bullet point here
- Second bullet point here
- Third bullet point here