

Model Selection

INSTRUCTOR: HONGJIE CHEN
MAY 24TH 2022

Which Model to Use

- Model 1: Training accuracy: 0.99
- Model 2: Training accuracy: 0.60
- Model 3: Training accuracy: 0.20

Outlook	Temperature	Humidity	Windy	PlayTennis
Sunny	Hot	High	False	No
Sunny	Hot	High	True	No
Overcast	Hot	High	False	Yes
Rainy	Mild	High	False	Yes
Rainy	Cool	Normal	False	Yes
Rainy	Cool	Normal	True	No
Overcast	Cool	Normal	True	Yes
Sunny	Mild	High	False	No
Sunny	Cool	Normal	False	Yes
Rainy	Mild	Normal	False	Yes
Sunny	Mild	Normal	True	Yes
Overcast	Mild	High	True	Yes
Overcast	Hot	Normal	False	Yes
Rainy	Mild	High	True	No

Case 1

- Recall validation dataset, (fake to be unseen)
- Model 1: Training accuracy: 0.99, validation accuracy: 0.24
- Model 2: Training accuracy: 0.60, validation accuracy: 0.58
- Model 3: Training accuracy: 0.20, validation accuracy: 0.21

Case 2

Overfitting v.s. Underfitting

What's an animal?

- Has (hairs=34752,eye=24.2mm,...) or (hairs=123242,eye=3.8mm,...) or...
- Overfitting (Too specific)
 - Poor generalization
 - High dimensional
 - Weakly regularized
 - Insufficient data
- Anything that moves
- Underfitting (Too general)
 - Low dimensional
 - Heavily regularized

Usually overfitting is our concern

Example of Different Fitting

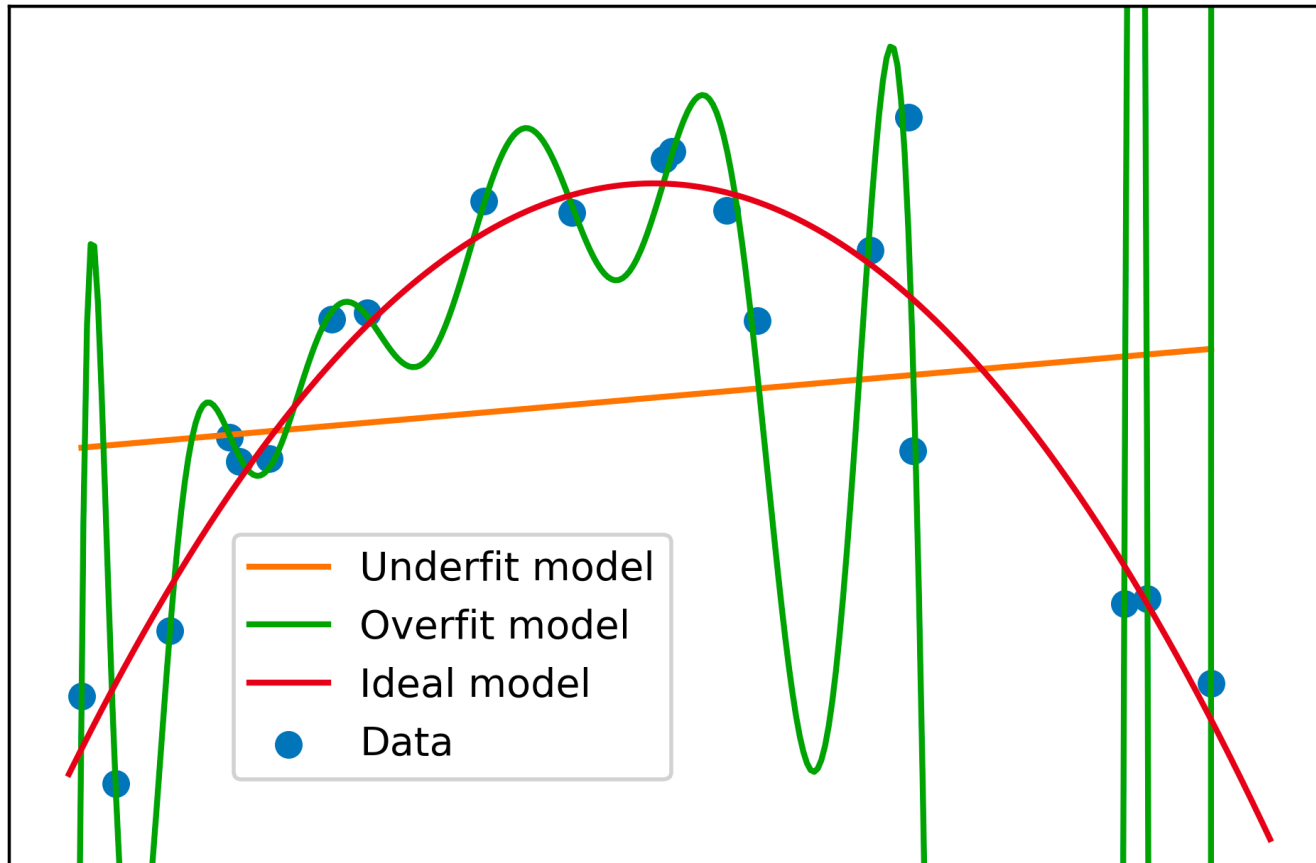
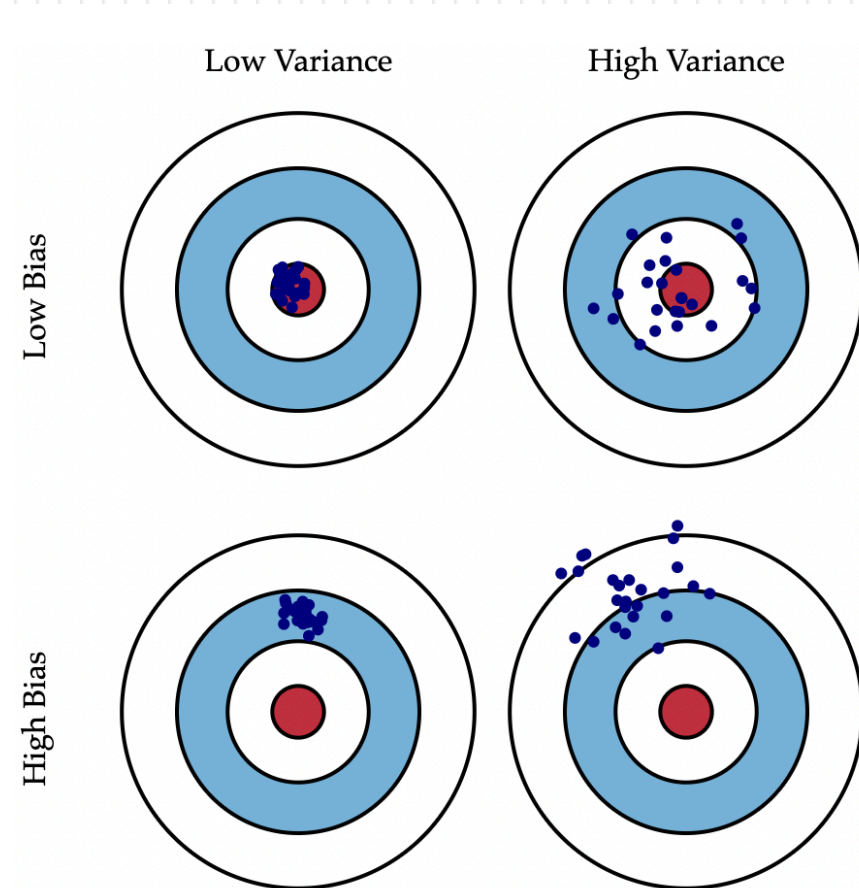


Figure credit: [link](#)

Bias v.s. Variance

- Four guns in a shooting game
- Both cause error
 - Bias is systematic
 - Variance is from random noise
- Reading: [statistical analysis](#)



Hold-out Validation Data Setting

- Use validation to avoid overfitting
- How to select validation dataset?

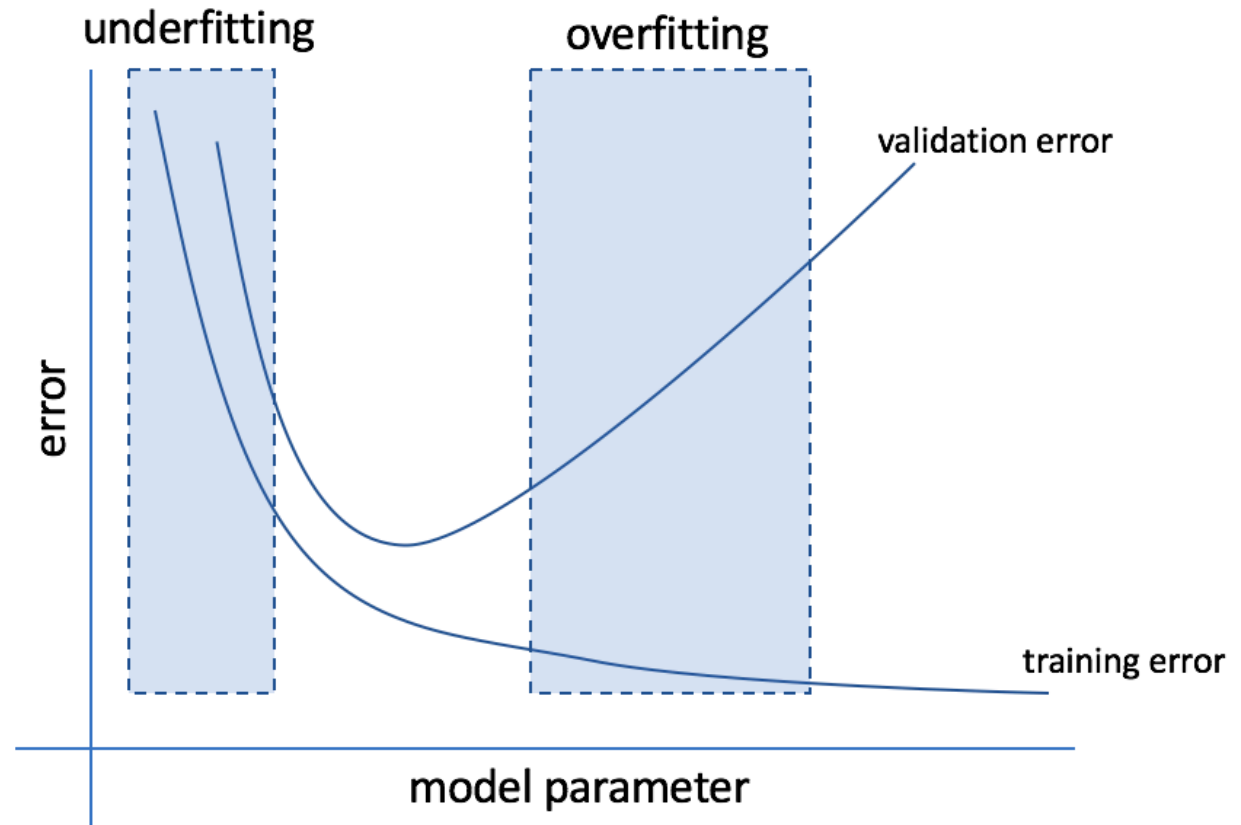


Figure credit: [link](#)

K-fold and Leave-one-out Cross Validation

- Split into K folds
 - One fold for validation
 - Others for training
- Select another fold for validation
 - Others for training
- And so on.
- Error is the average error of the K models
- Leave-one-out: one fold contains only one sample

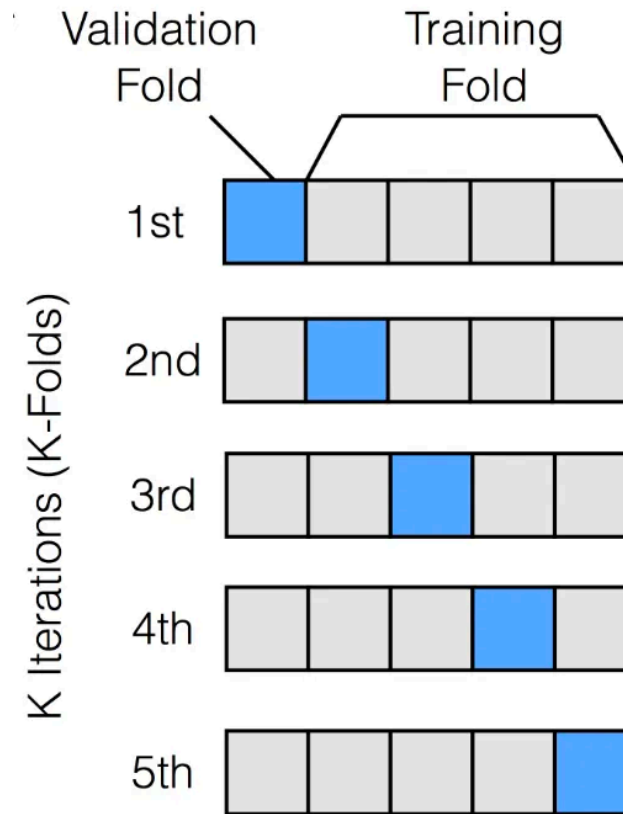


Figure credit: [link](#)

Common Practice in Experiments

- Testing data are hidden from training
 - [Kaggle competition](#)
- Use validation accuracy to fine tune hyperparameter

- Evaluate on testing data