

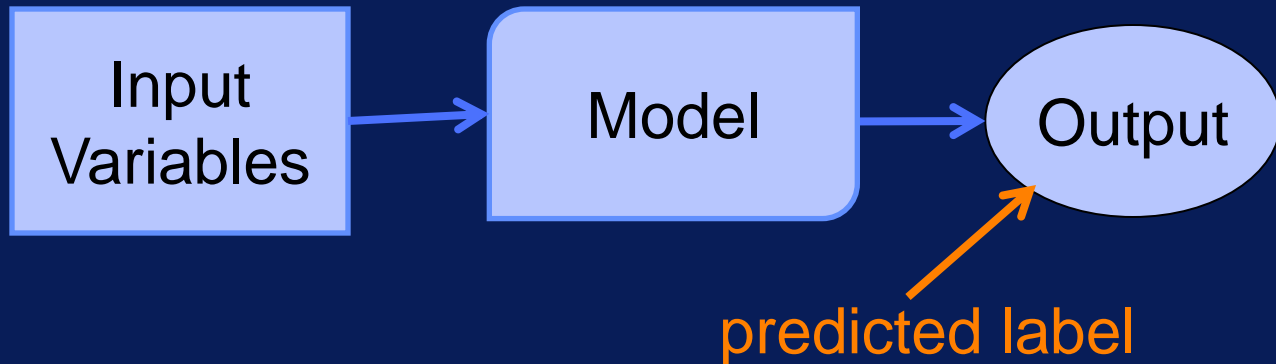
Generalization & Overfitting

After this video you will be able to..

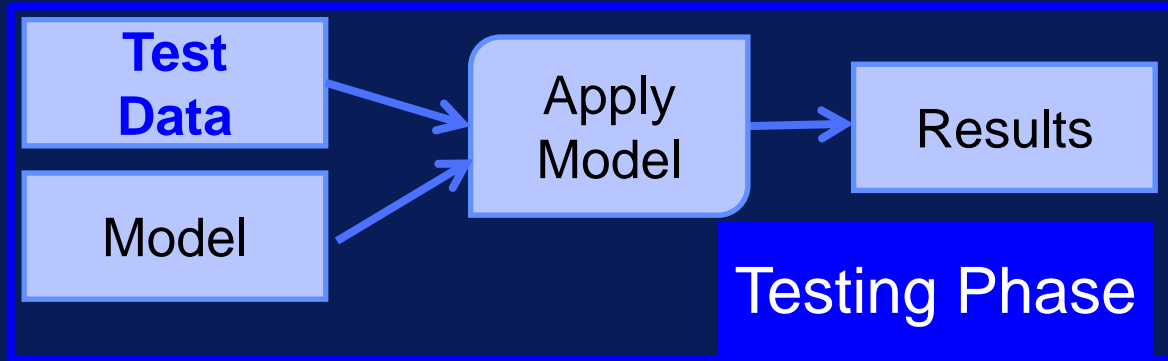
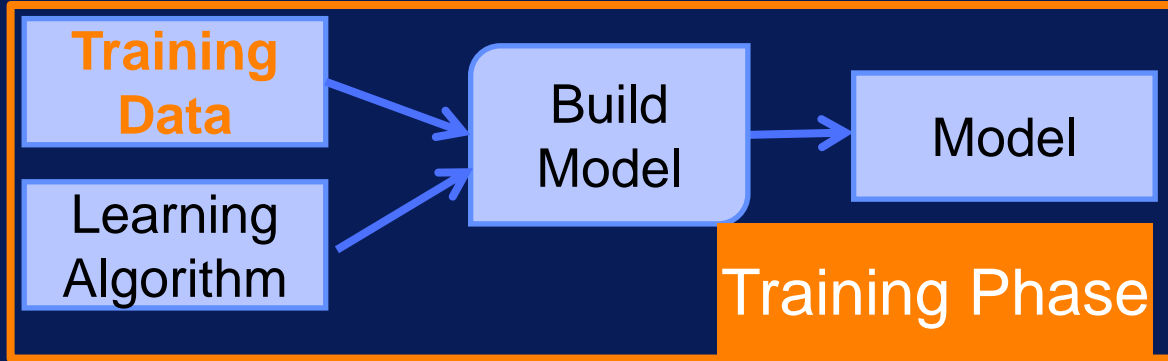
- Define what generalization is
- Describe how overfitting is related to generalization
- Explain why overfitting should be avoided

Errors in Classification

- Success: Output = Target ← true label
- Error: Output != Target
- Error rate = Error = Misclassification Error
 - # errors / # samples = % error



Training vs. Testing Phases



Errors in Classification

Error on
Training
Data

Training Error

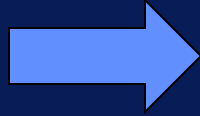
Error on
Test
Data

Test Error

Test error indicates how well
model will perform on new data!

Generalization

*Performs well
on new data*

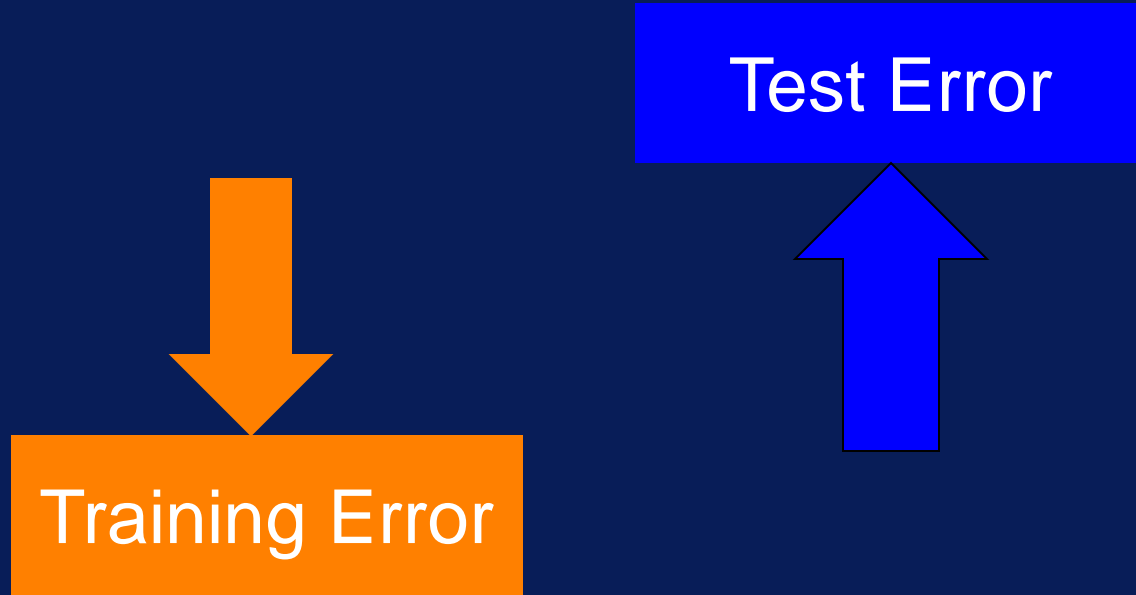


*Good
Generalization*



Test Error = Generalization Error

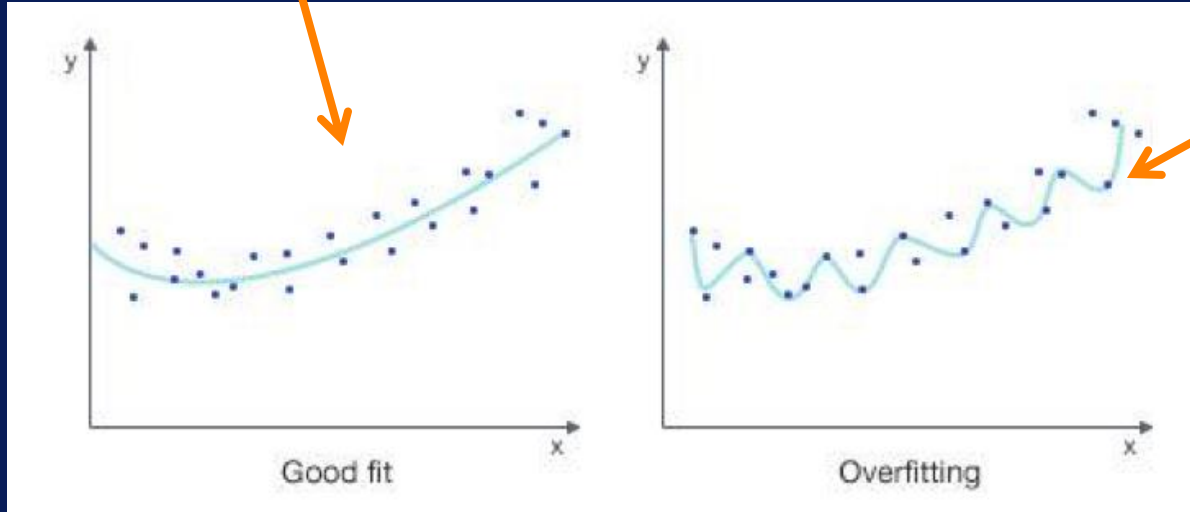
Overfitting



Overfitting

Model is fitting to structure of data

Model is fitting to noise in data

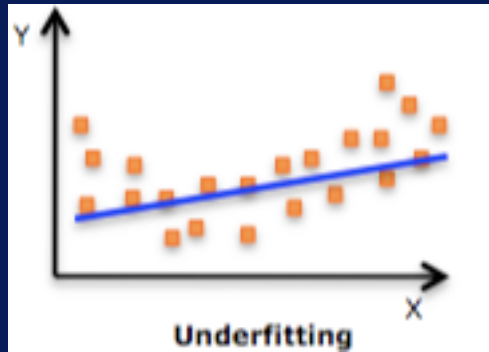


Source: <http://blog.fliptop.com/blog/2015/03/02/bias-variance-and-overfitting-machine-learning-overview/>

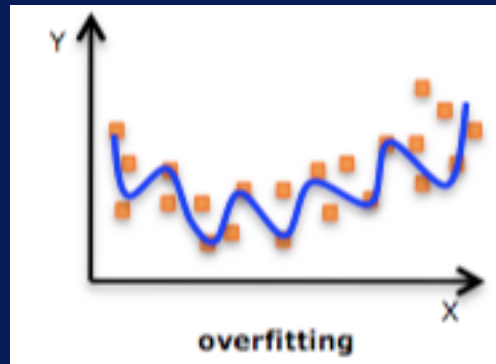
Overfitting & Generalization

Overfitting → Poor
Generalization

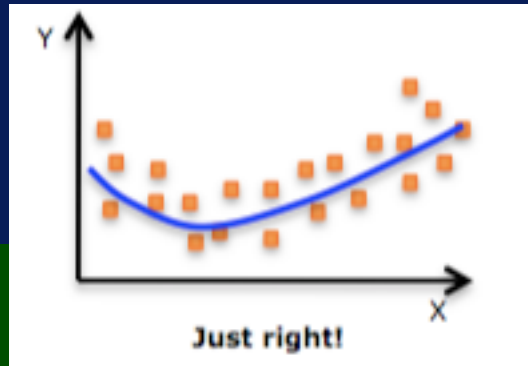
Overfitting & Underfitting



Underfitting

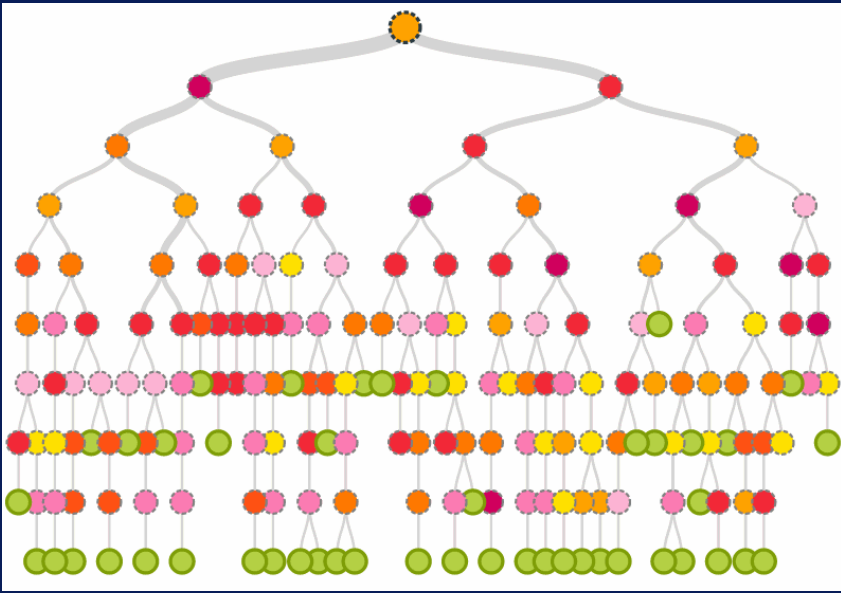


Overfitting



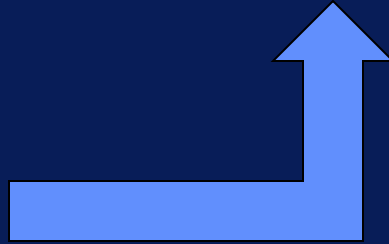
Just Right

What Causes Overfitting?

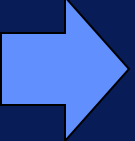


Overly complex model

Overfitting



Generalization & Overfitting

Overfitting  Good
Generalization