

# DATAPHILLY



Machine Learning and Performance Evaluation

Sebastian Raschka

Wednesday, November 30, 2016

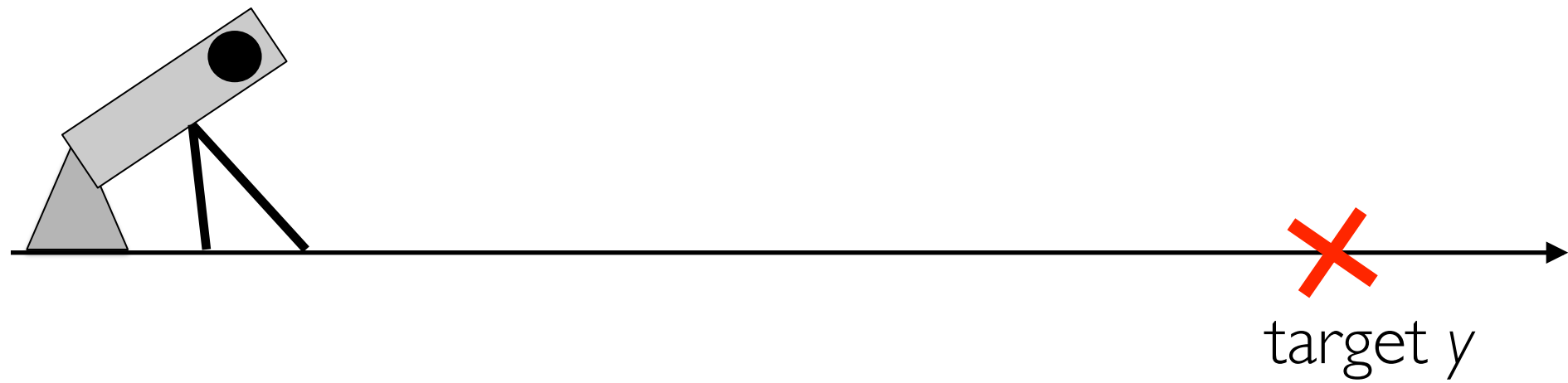
6:30 PM to 7:00 PM

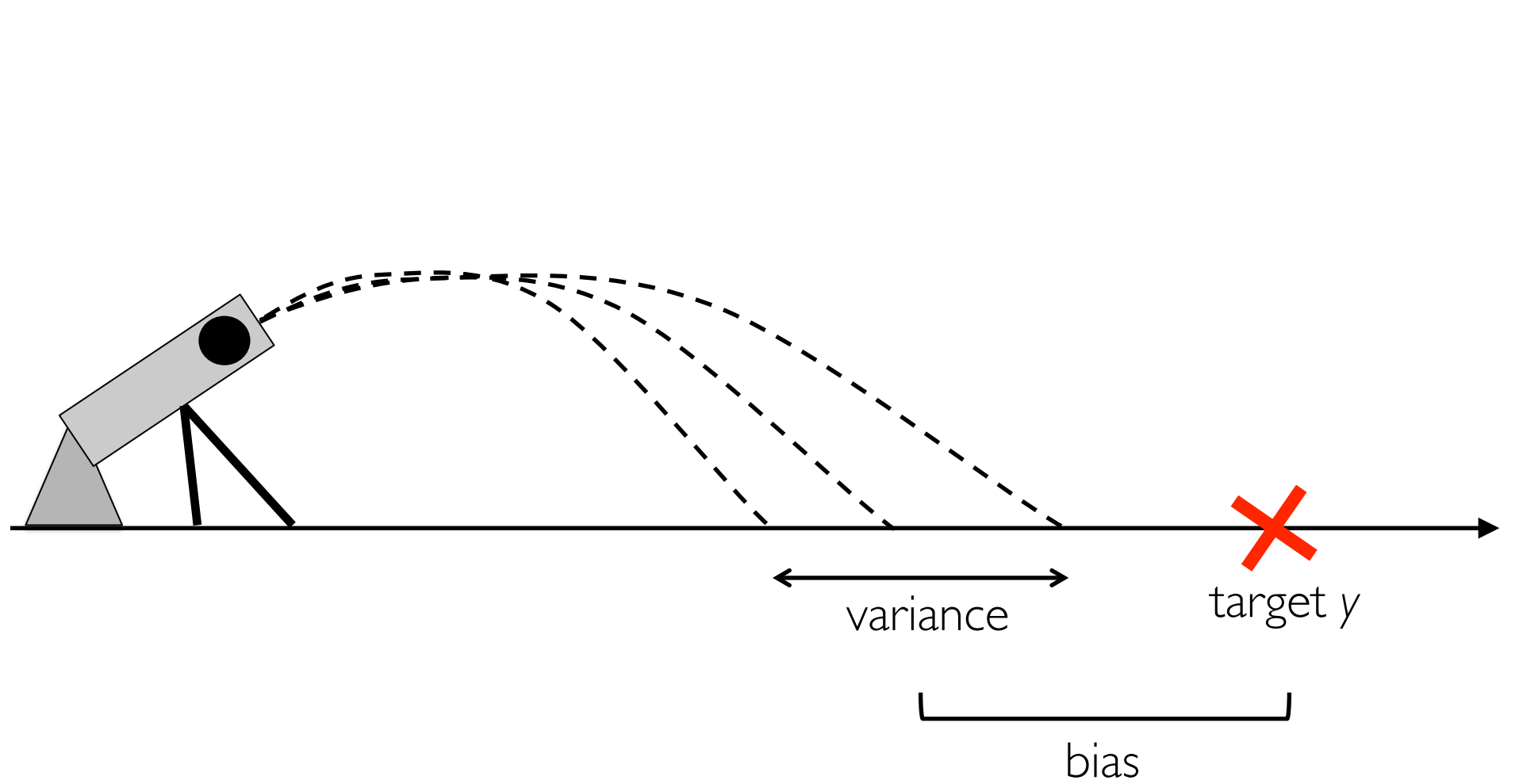
441 N 5th St, Suite 301, Philadelphia, PA

# Estimating the Performance of Predictive Models

Why bother?

- ① Generalization Performance
- ② Model Selection
- ③ Algorithm Selection

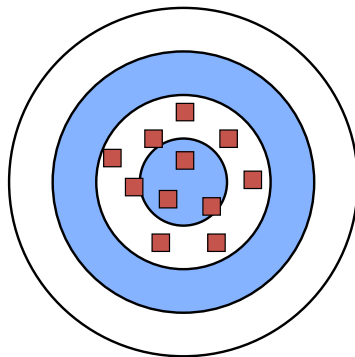
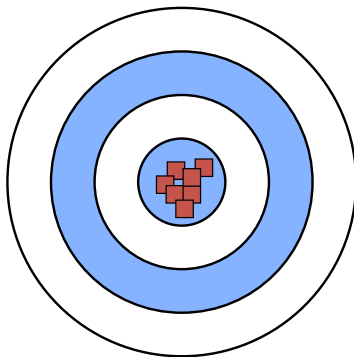




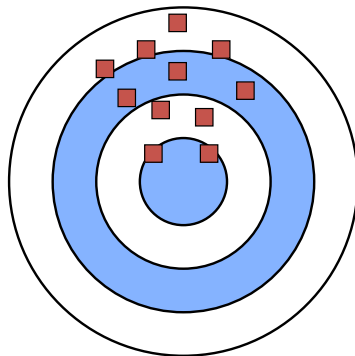
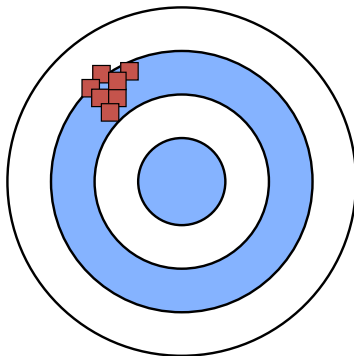
**Low Variance**  
(Precise)

**High Variance**  
(Not Precise)

**Low Bias**  
(Accurate)



**High Bias**  
(Not Accurate)



expected estimated value

$$\text{BIAS} = E[\hat{\beta}] - \beta$$

$$\text{VARIANCE} = E\left[(\hat{\beta} - E[\hat{\beta}])^2\right]$$

# Performance Estimates

- Absolute vs Relative

Ranking Tasks

- ① Generalization Performance
- ② Model Selection
- ③ Algorithm Selection



TRAIN

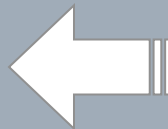


TRAIN

TEST

Sources of Bias and Variance

TRAIN



TEST

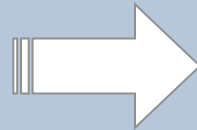
Pessimistic Bias 





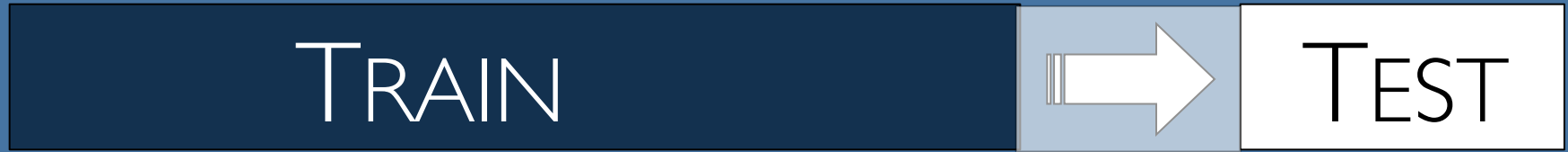
\* SoftMax Classifier on a small MNIST subset

TRAIN

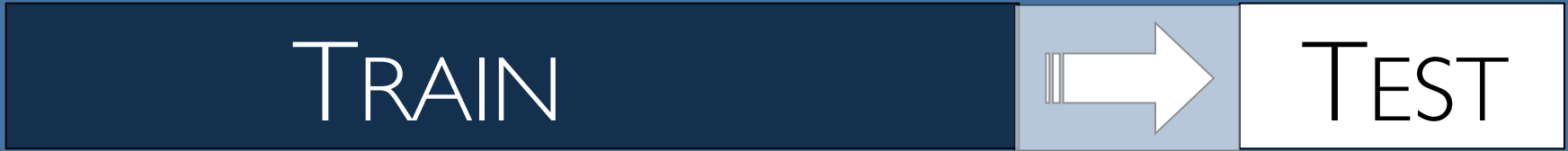


TEST

Pessimistic Bias 



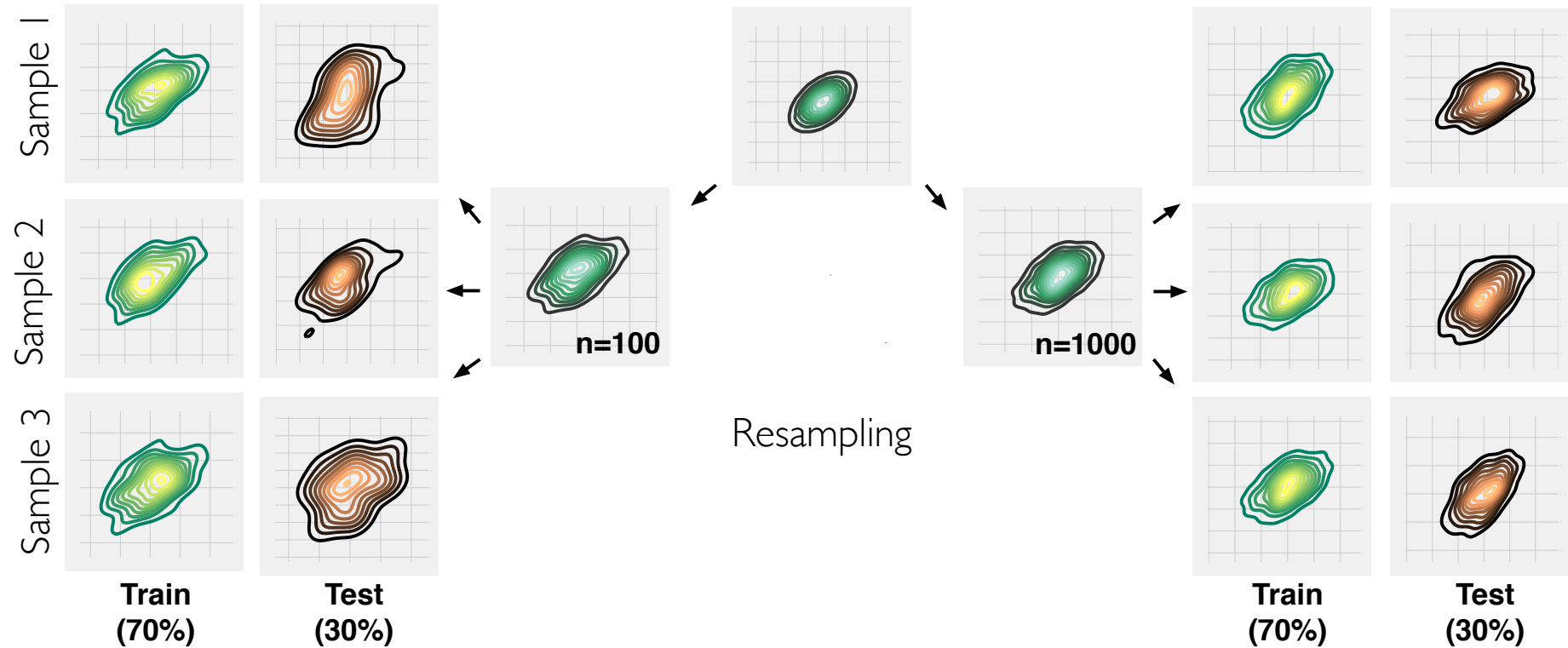
Pessimistic Bias



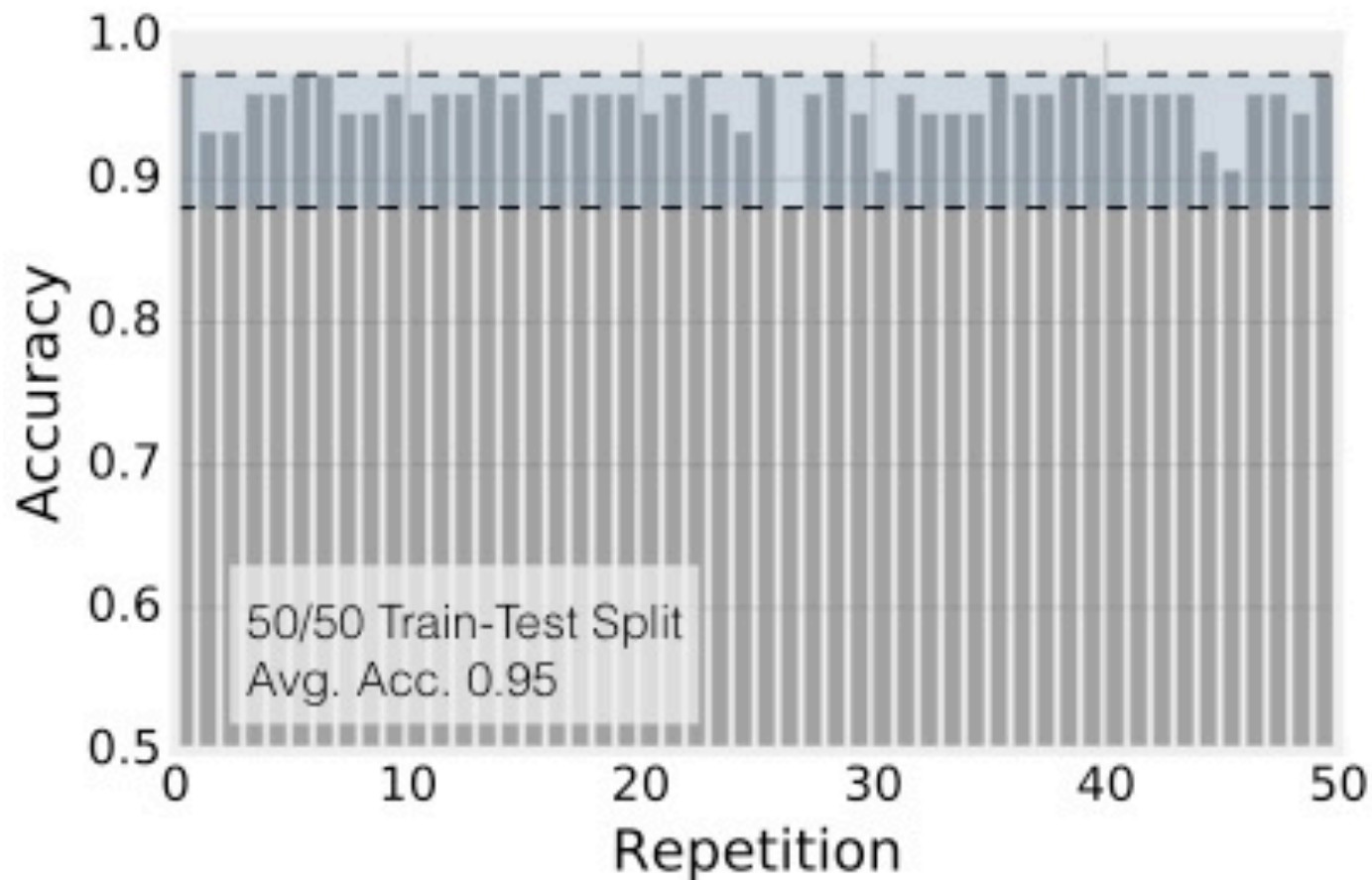
Variance



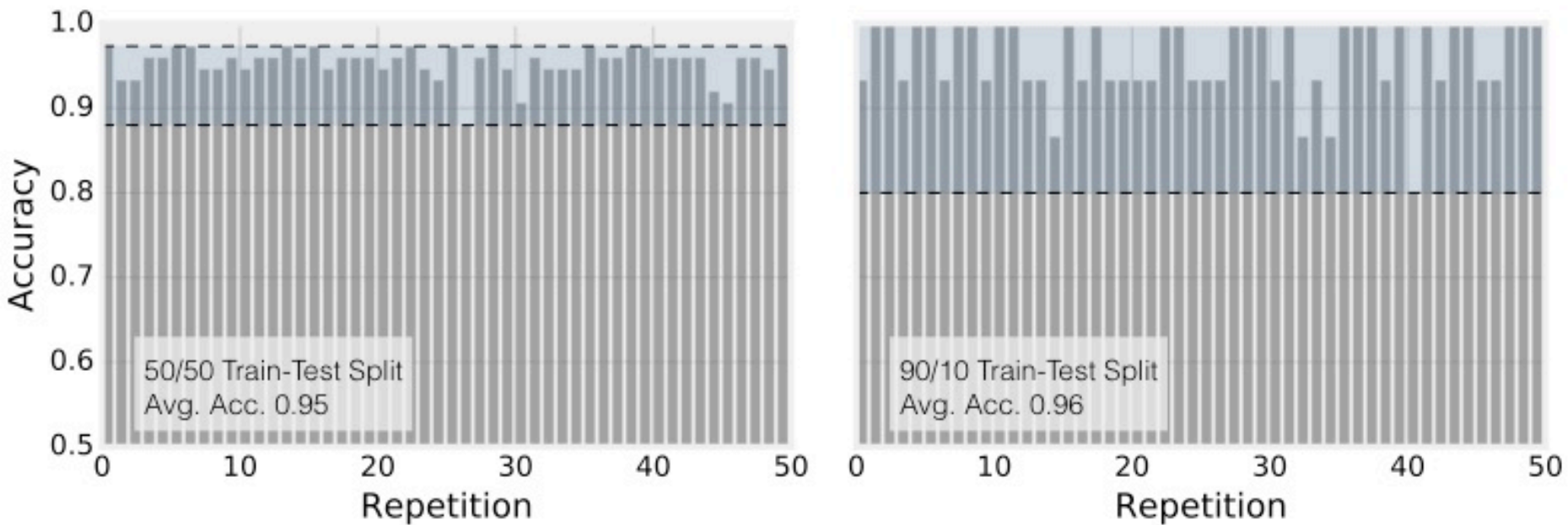
Real World Distribution





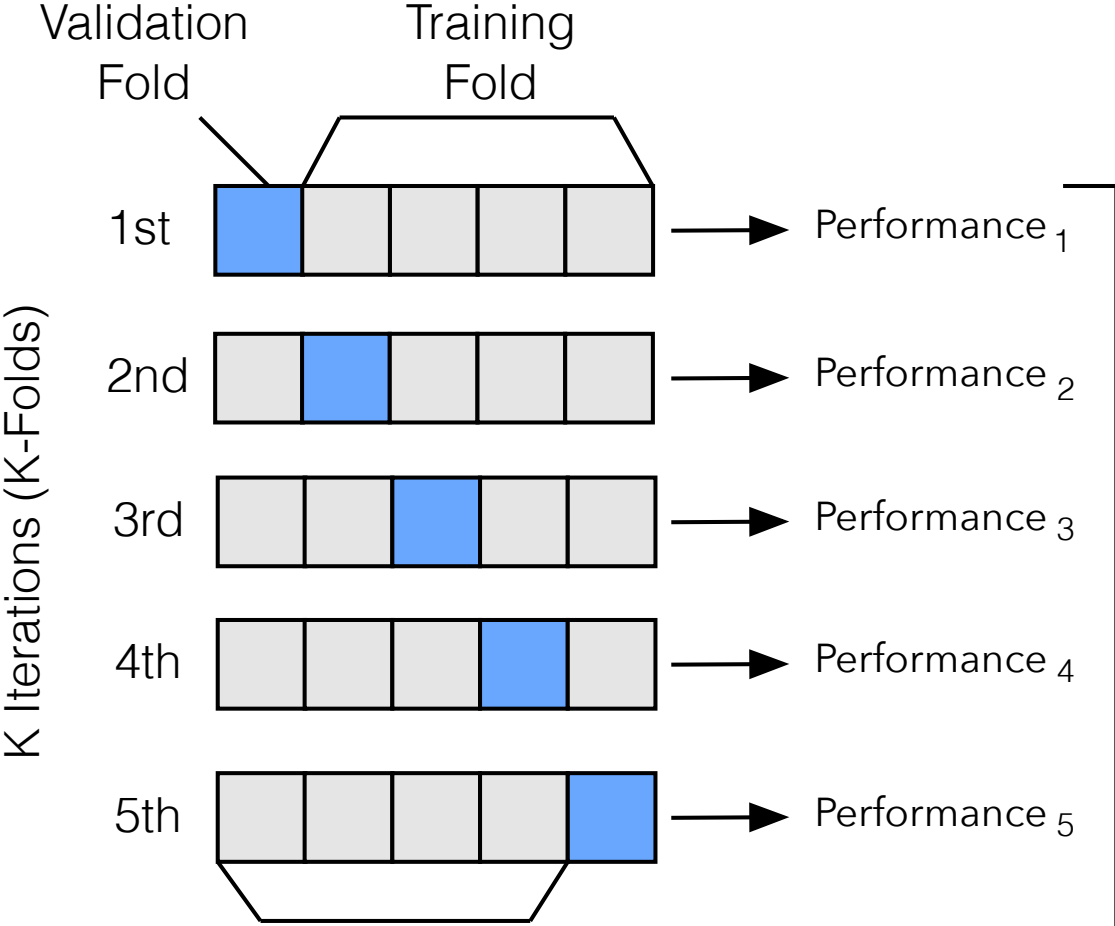


\* 3-NN on Iris dataset



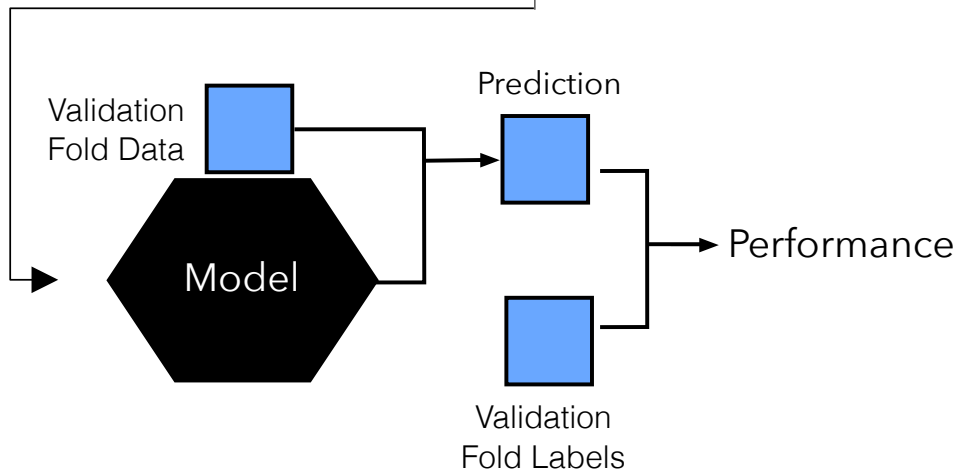
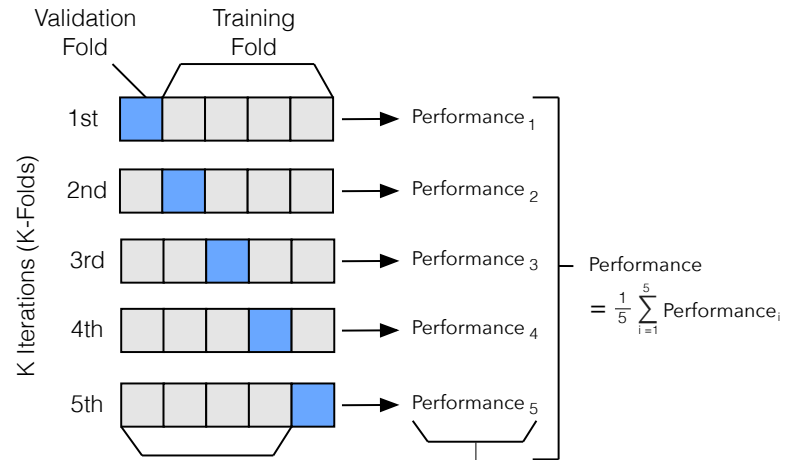
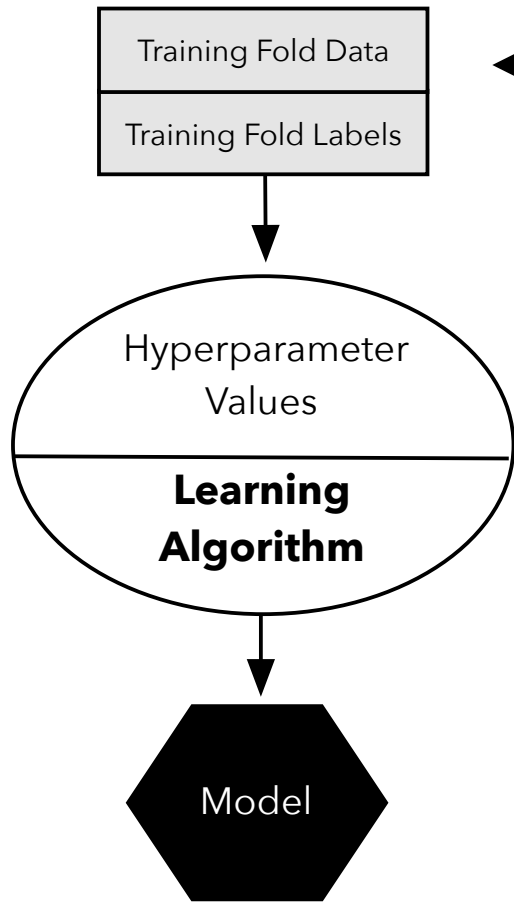
\* 3-NN on Iris dataset

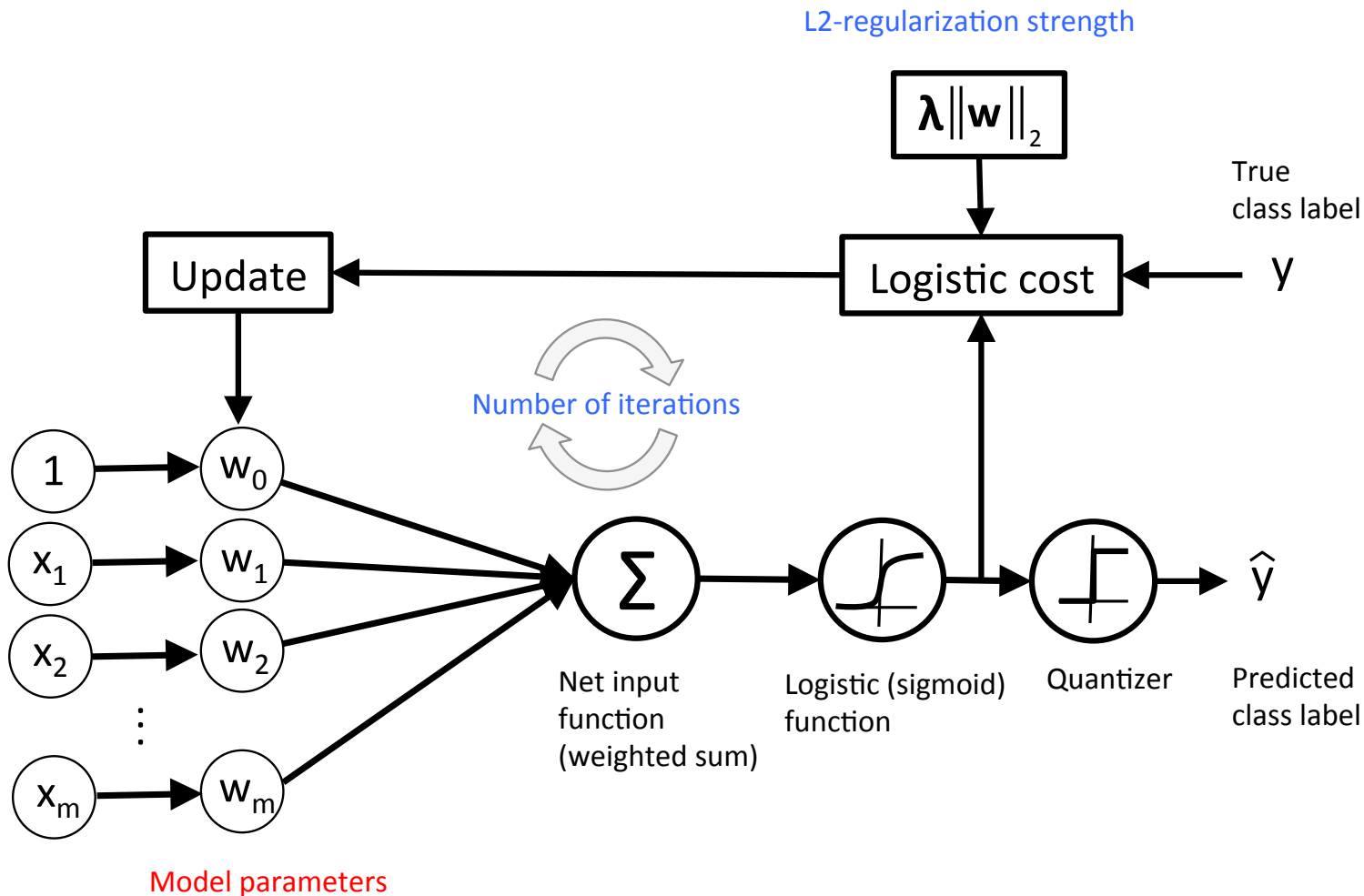
# K-fold Cross Validation



Performance

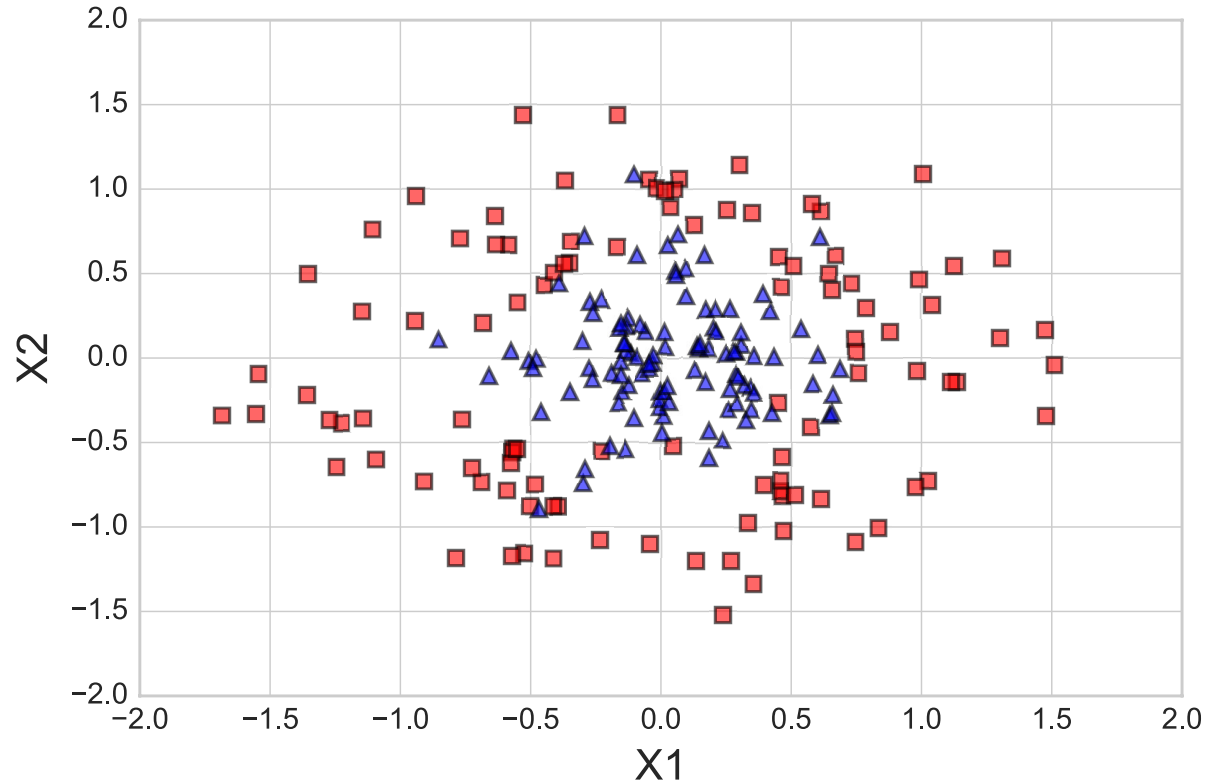
$$= \frac{1}{5} \sum_{i=1}^5 \text{Performance}_i$$





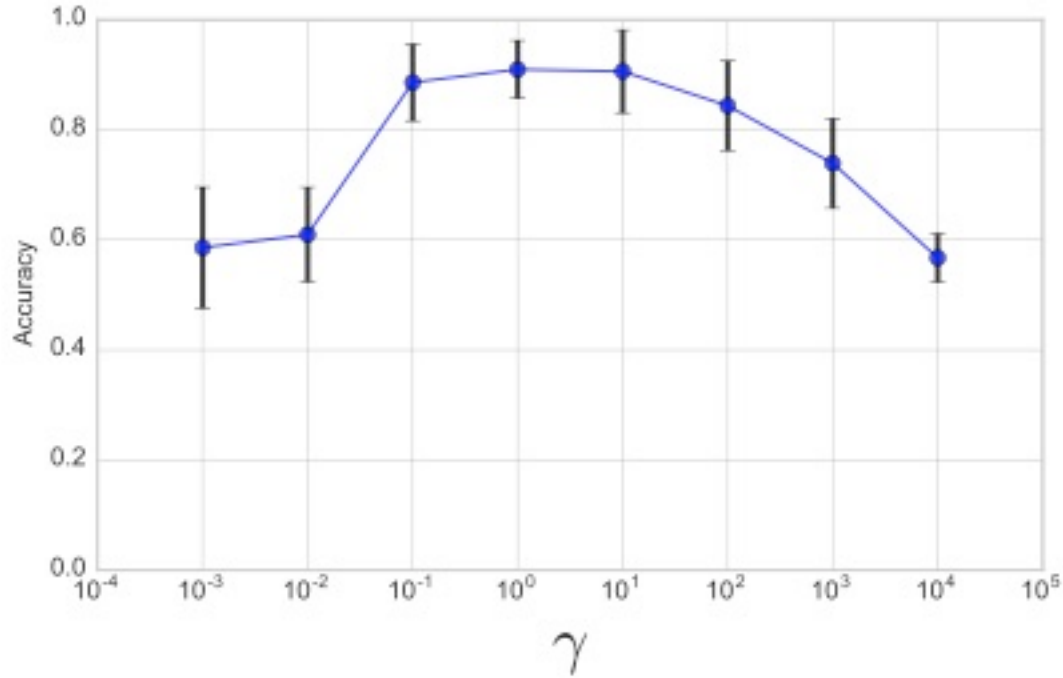
# The law of parsimony

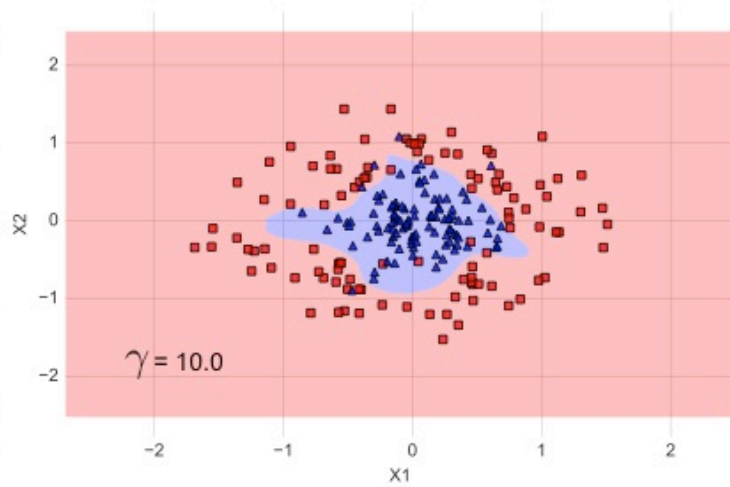
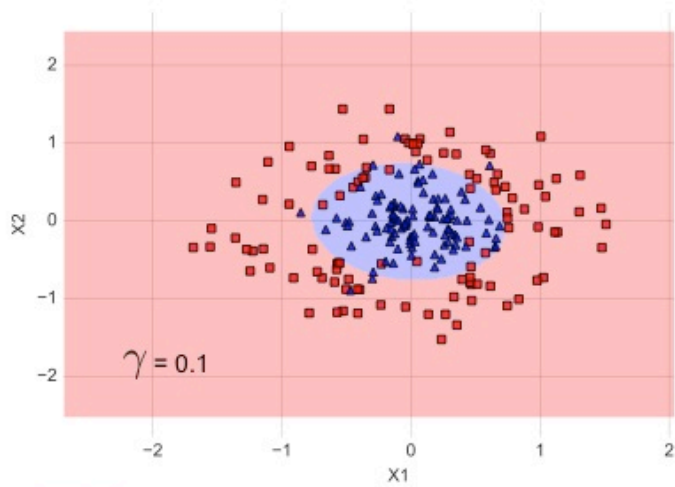
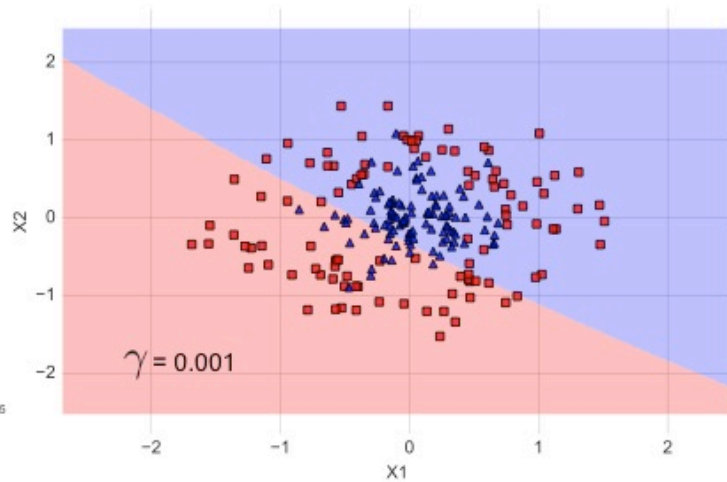
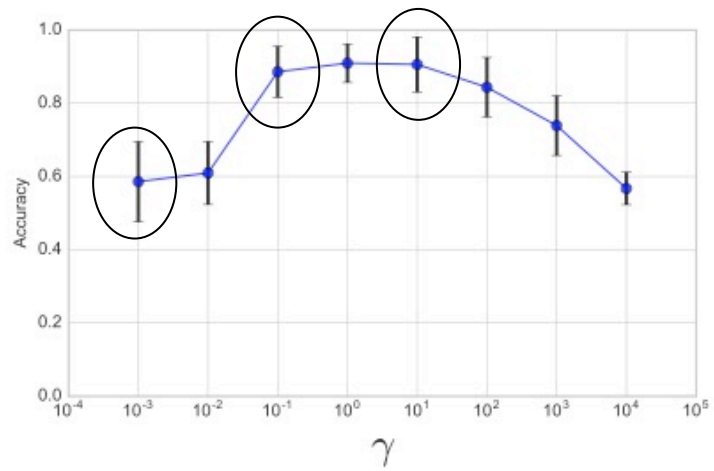
## I-standard error method



# The law of parsimony

## I-standard error method

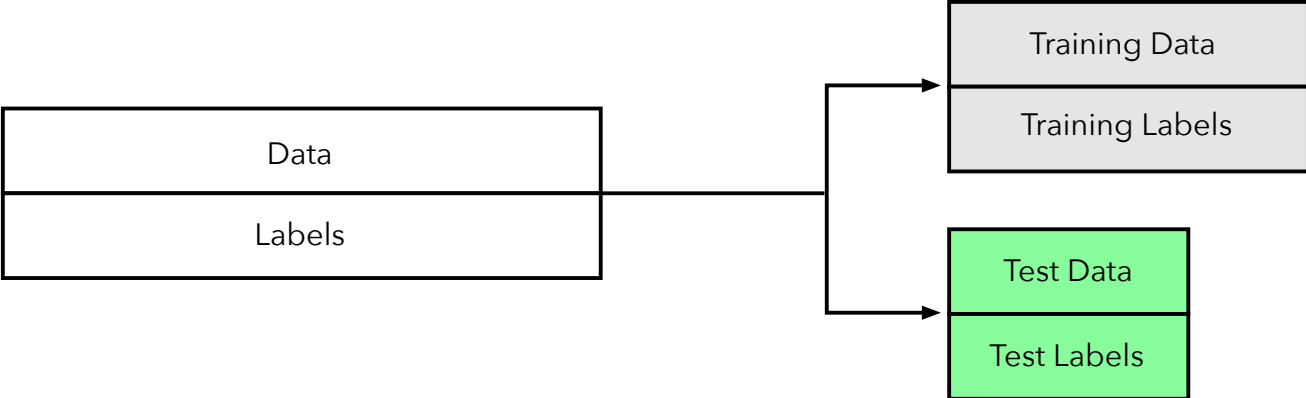




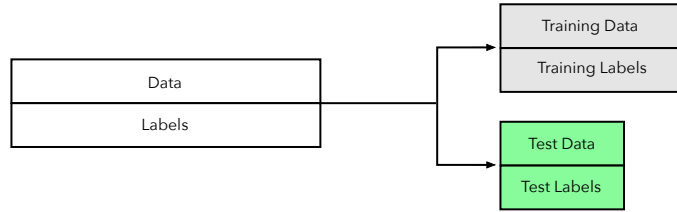


# K-fold for Model Selection step-by-step

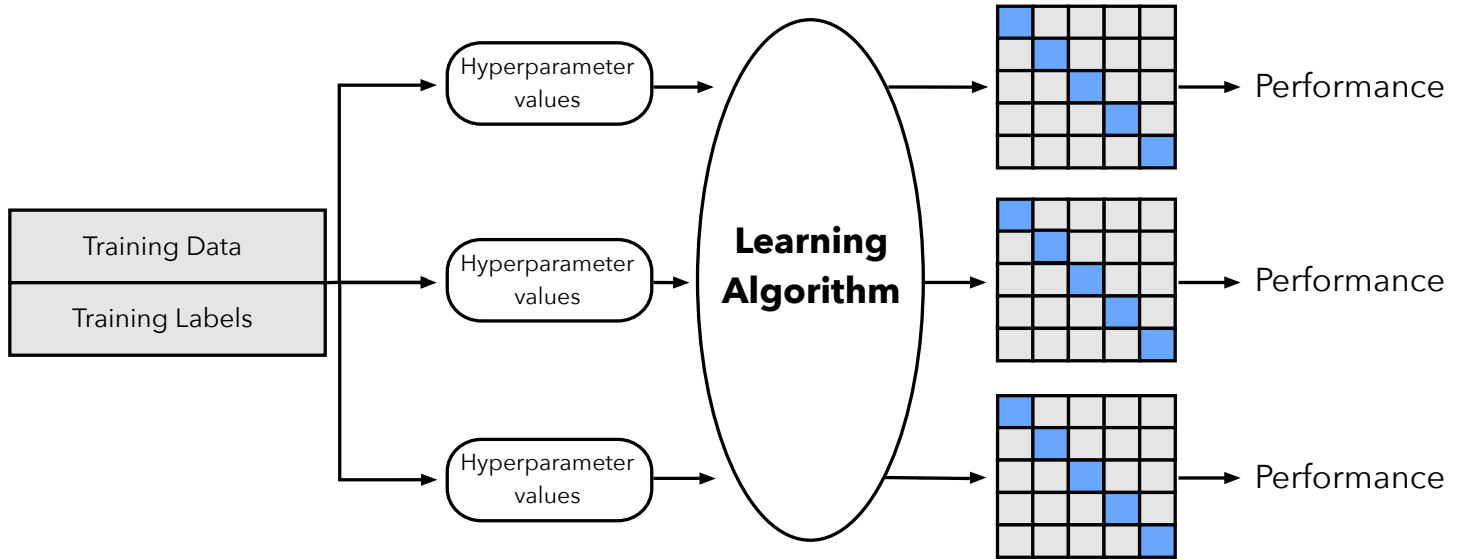
1



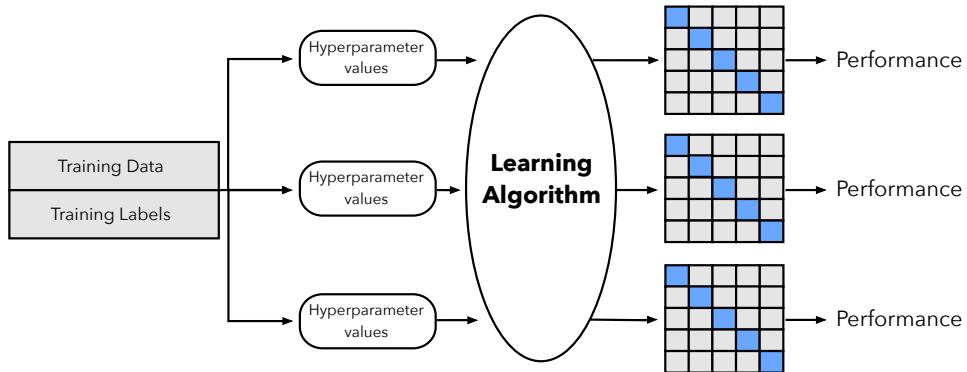
1



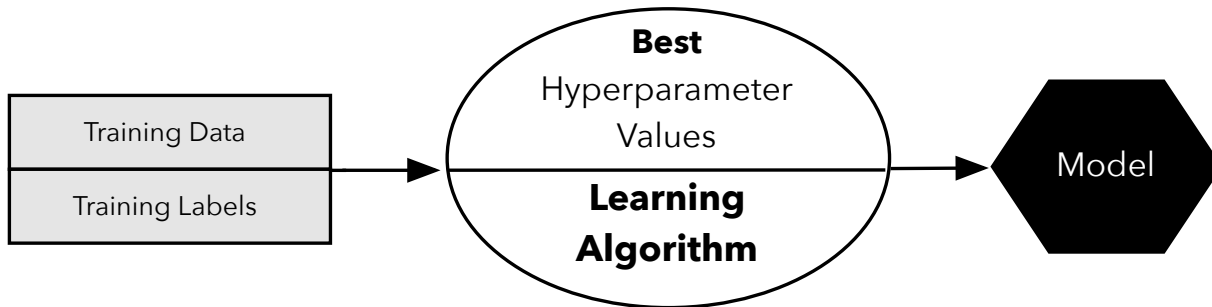
2



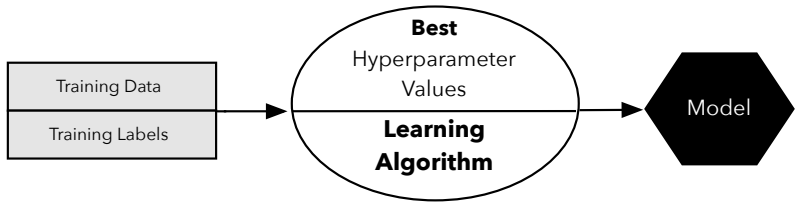
2



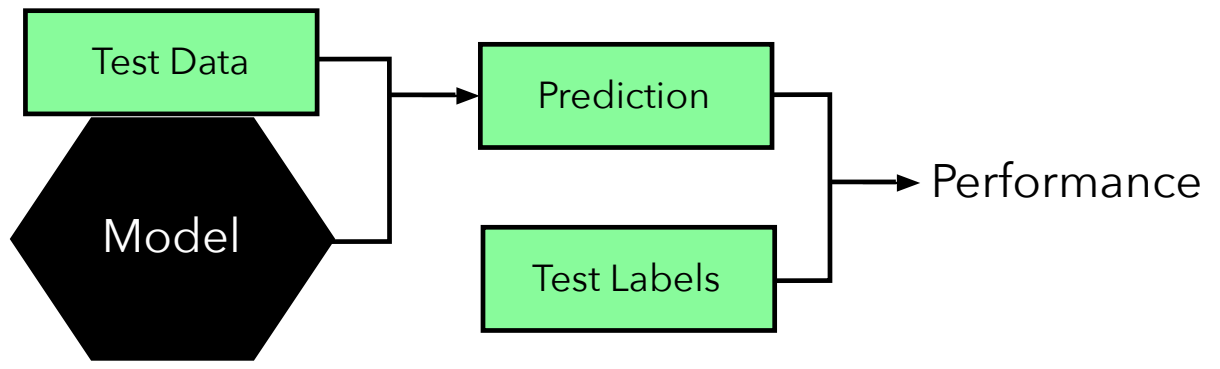
3



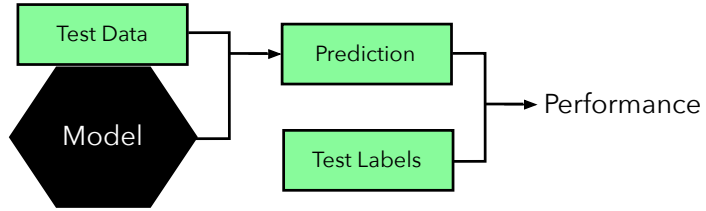
3



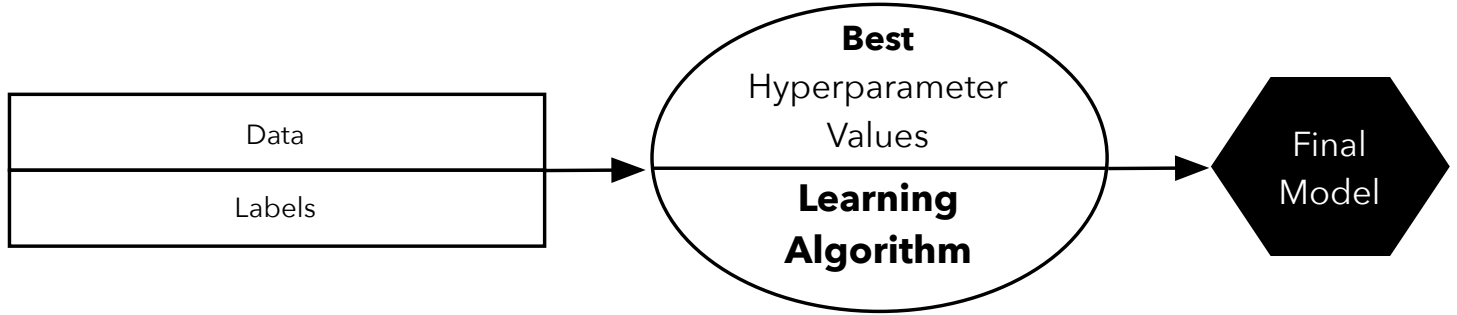
4



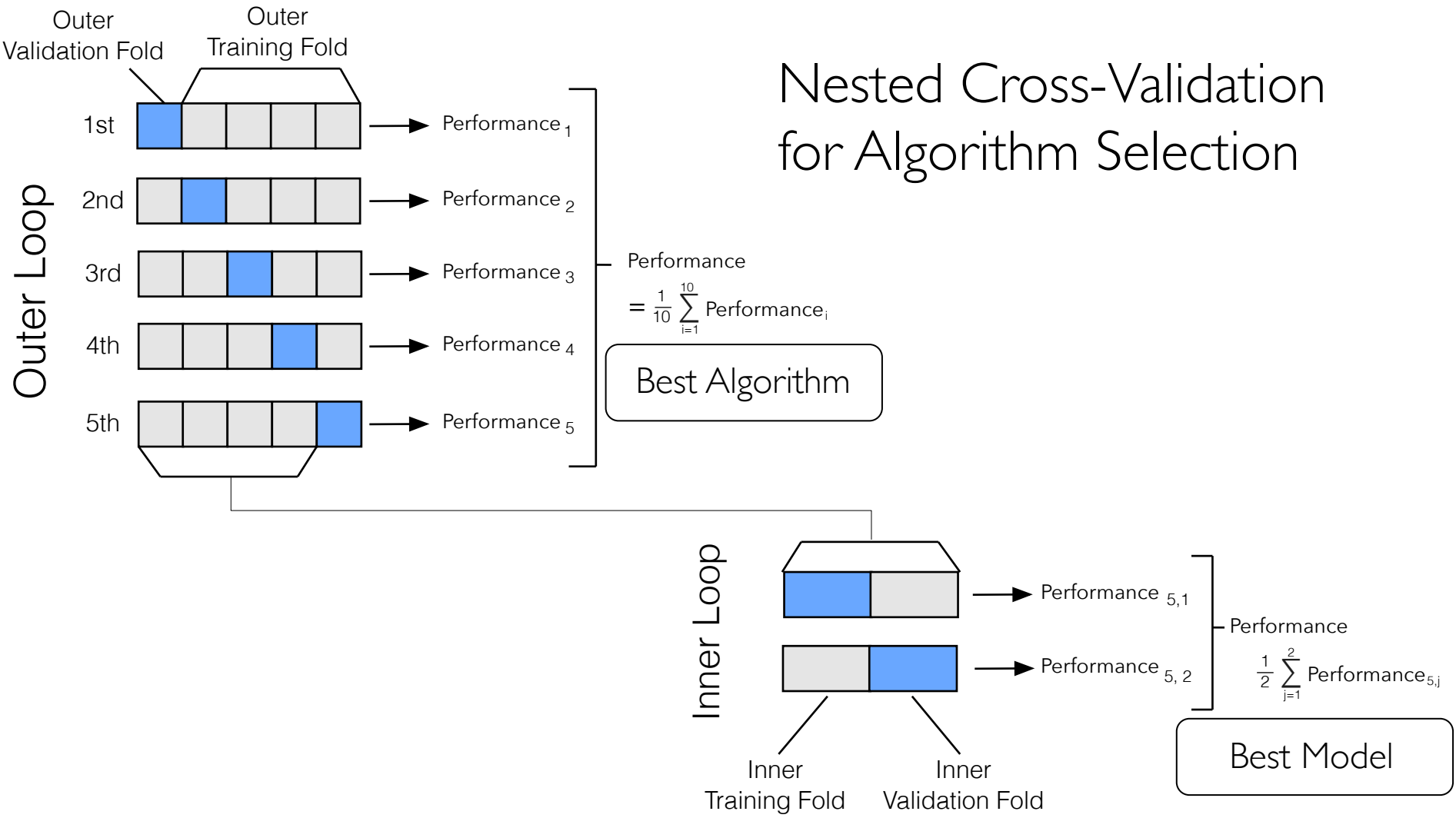
4



5



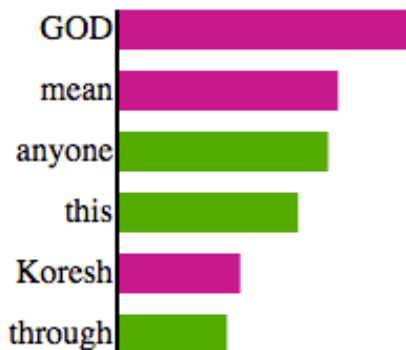
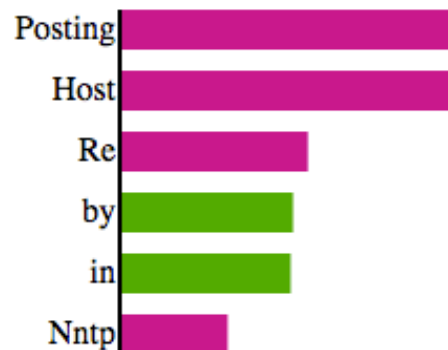
# Nested Cross-Validation for Algorithm Selection



# Beyond Performance Metrics

Ideal features that are ...

- discriminatory
- salient
- invariant

**Algorithm 1****Words that A1 considers important:****Predicted:** Atheism**Prediction correct:****Algorithm 2****Words that A2 considers important:****Predicted:** Atheism**Prediction correct:****Document**

From: pauld@verdix.com (Paul Durbin)  
 Subject: Re: DAVID CORESH IS! **GOD!**  
 Nntp-Posting-Host: sarge.hq.verdix.com  
 Organization: Verdix Corp  
 Lines: 8

**Document**

From: pauld@verdix.com (Paul Durbin)  
 Subject: **Re:** DAVID CORESH IS! GOD!  
**Nntp-Posting-Host:** sarge.hq.verdix.com  
 Organization: Verdix Corp  
 Lines: 8



THANK YOU!



<https://github.com/rasbt>



<http://sebastianraschka.com>  
[mail@sebastianraschka.com](mailto:mail@sebastianraschka.com)



[@rasbt](https://twitter.com/rasbt)