Compressed Sensing & Artificial Intelligence

M229 Advanced Topics in MRI

Kyung Sung, Ph.D.

2020.05.28

Class Business

- Project abstract due on 6/5 (Fri)
- Project presentation file on 6/11 (Thurs)
- Guest Lecturers:
  - Dr. Yingli Yang (6/2)
  - Dr. Fabien Scalzo (6/4)

Today's Topics

- Compressed sensing
  - Compressibility or sparsity
  - Incoherent measurement
  - Reconstruction
- Machine learning / artificial intelligence
  - Model evaluation
  - Model selection

Fast MRI Techniques

- Many reconstruction methods minimize aliasing artifacts by exploiting *information redundancy* (or *prior knowledge*)
  - Parallel imaging
  - Compressed sensing

Donoho, IEEE TIT, 2006
Candes et al., Inverse Problems, 2007

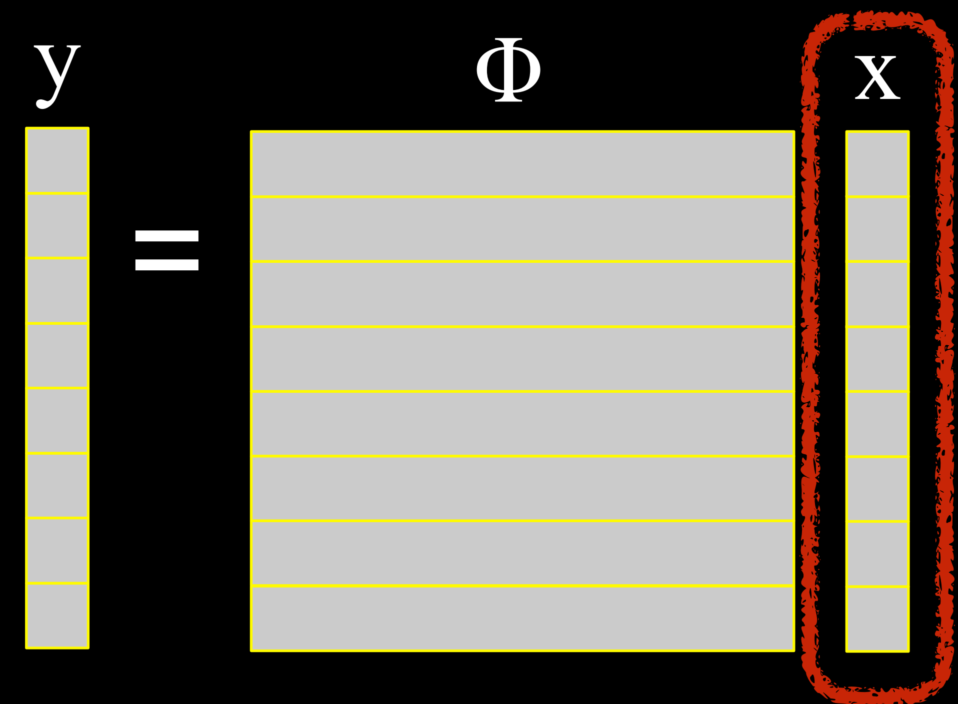
What is Compressed Sensing?

- CS is about acquiring a sparse signal in a most efficient way (subsampling) with the help of an incoherent projecting basis

What is Compressed Sensing?

- CS is about acquiring a sparse signal in a most efficient way (subsampling) with the help of an incoherent projecting basis

8 Equations
8 Unknowns

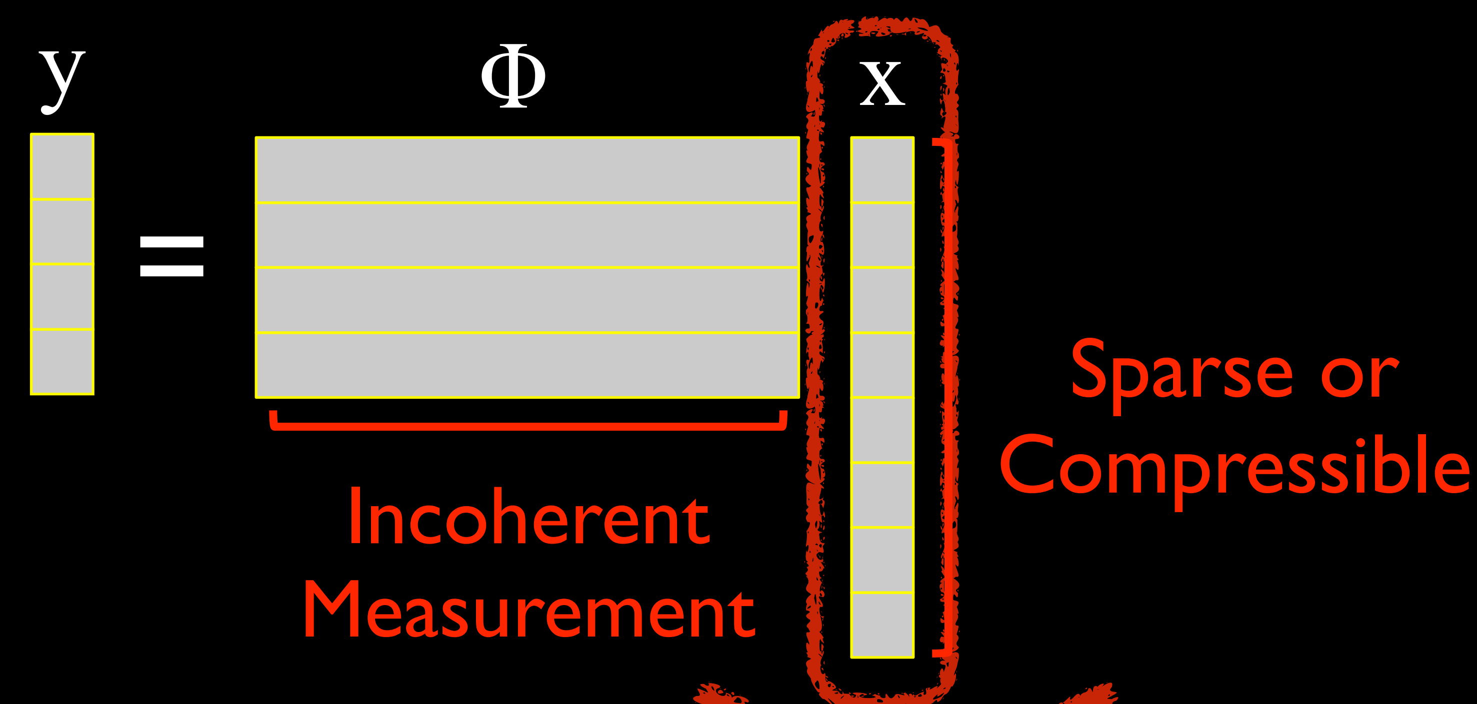


$$x = \Phi^{-1} y$$

What is Compressed Sensing?

- CS is about acquiring a sparse signal in a most efficient way (subsampling) with the help of an incoherent projecting basis

4 Equations
8 Unknowns

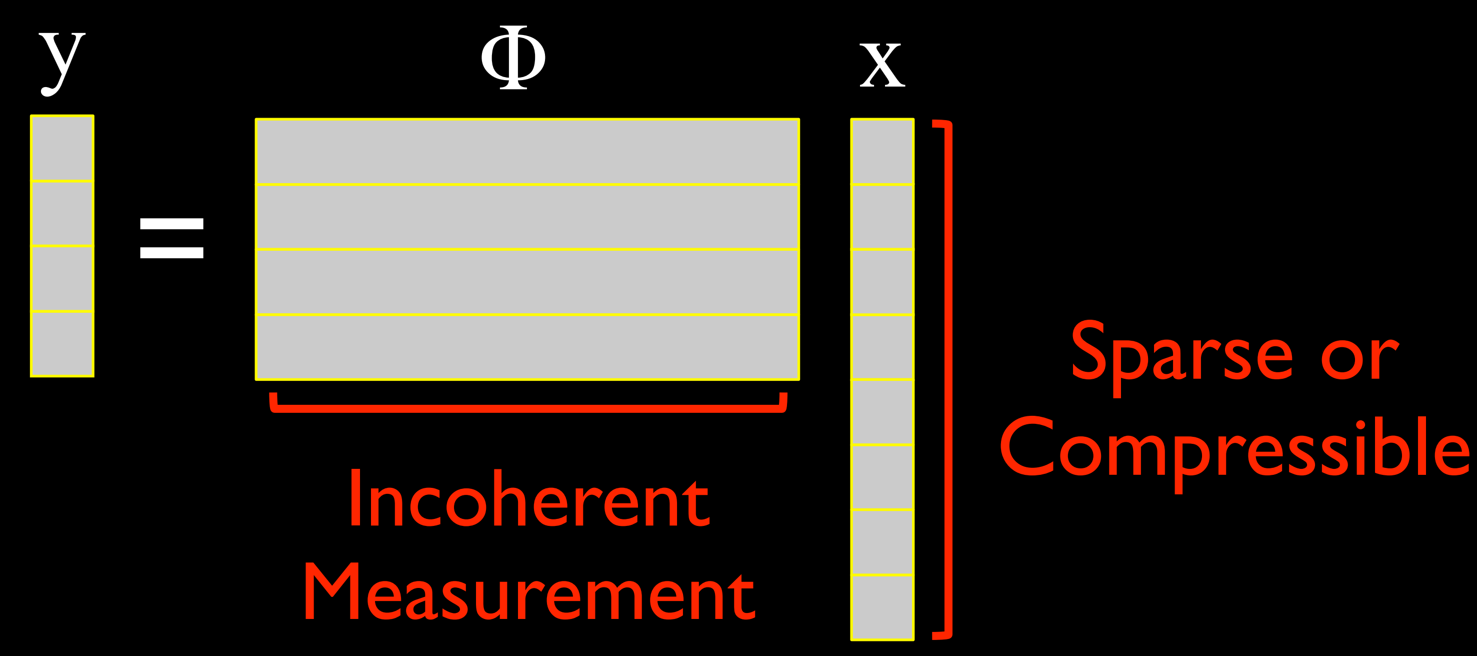


$\cancel{x = \Phi^{-1} y}$

What is Compressed Sensing?

- CS is about acquiring a sparse signal in a most efficient way (subsampling) with the help of an incoherent projecting basis

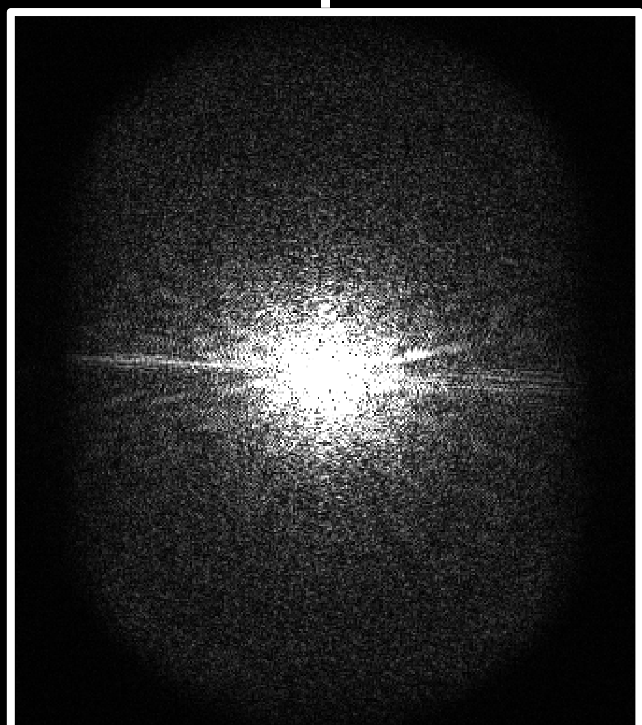
4 Equations
8 Unknowns



We still can find 8 unknowns!

Compressed Sensing MRI

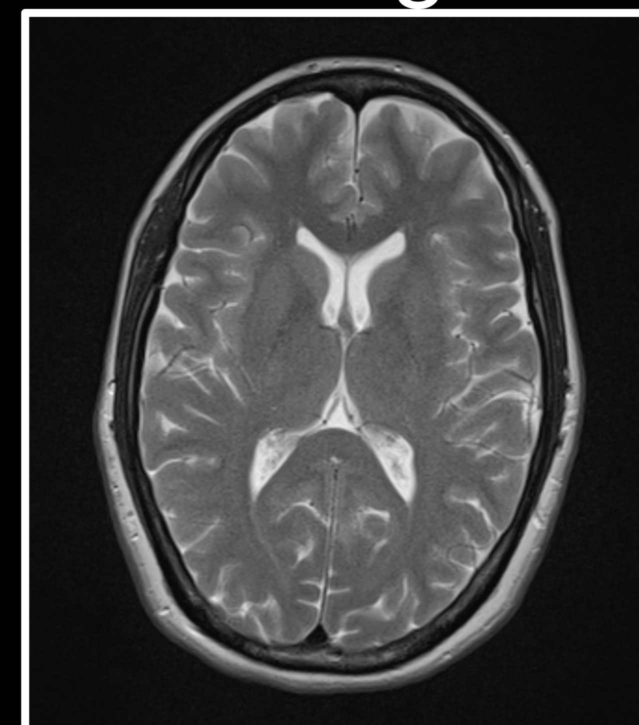
k-space



Inverse Fourier Transform Φ^{-1}

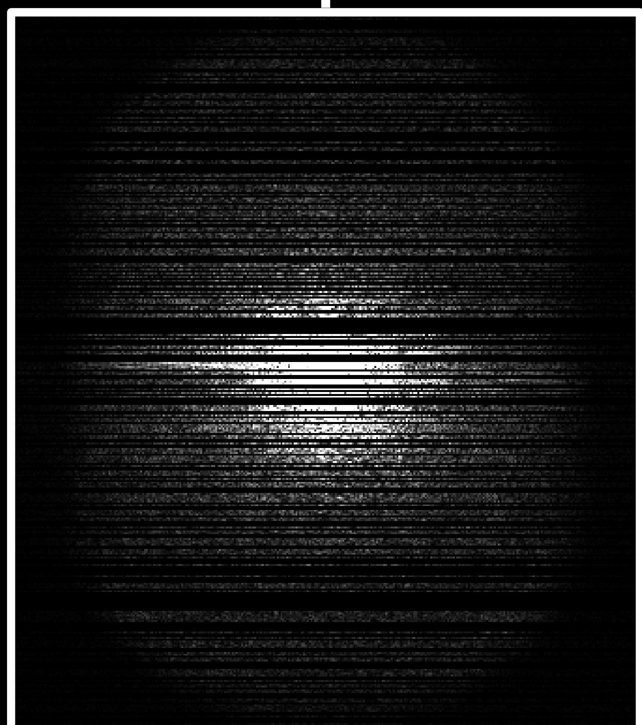
$$x = \Phi^{-1}y$$

Image



Compressed Sensing MRI

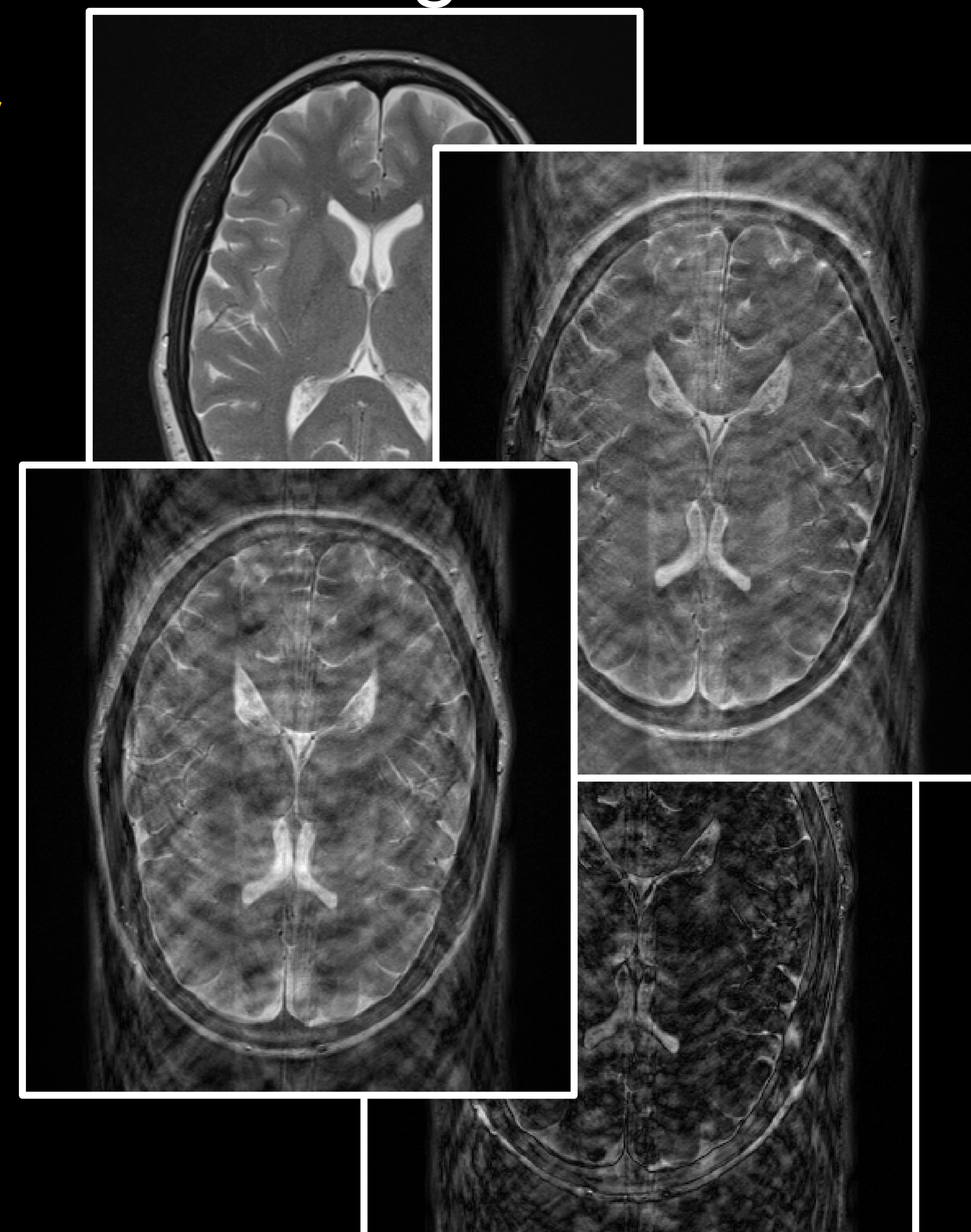
k-space



~~Inverse Fourier Transform~~ $\cancel{\Phi^{-1}}$

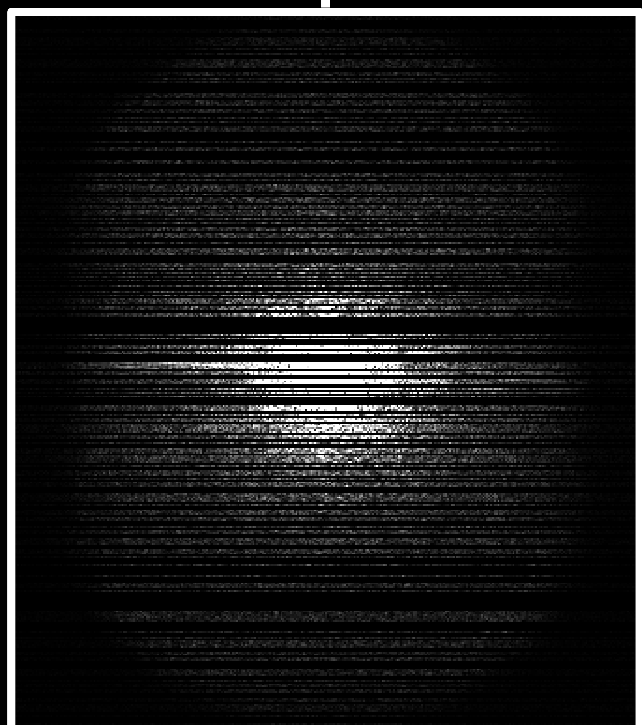
$$\cancel{x = \Phi^{-1} y}$$

Image



Compressed Sensing MRI

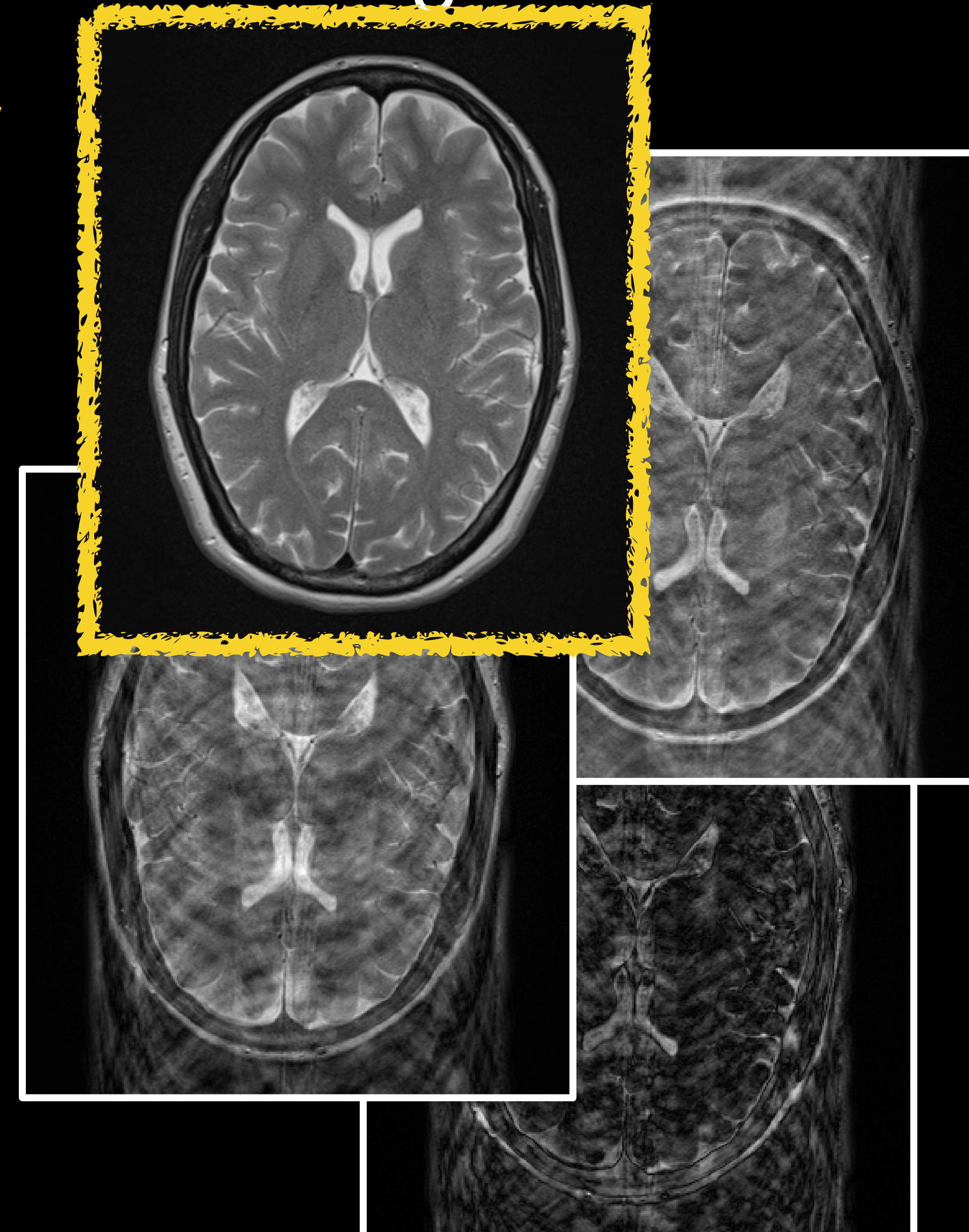
k-space



~~Inverse Fourier Transform~~ $\cancel{\Phi^{-1}}$

$$\cancel{x = \Phi^{-1} y}$$

Image



Choose the most compressible image matching data (*systematic optimization*)

Math Background

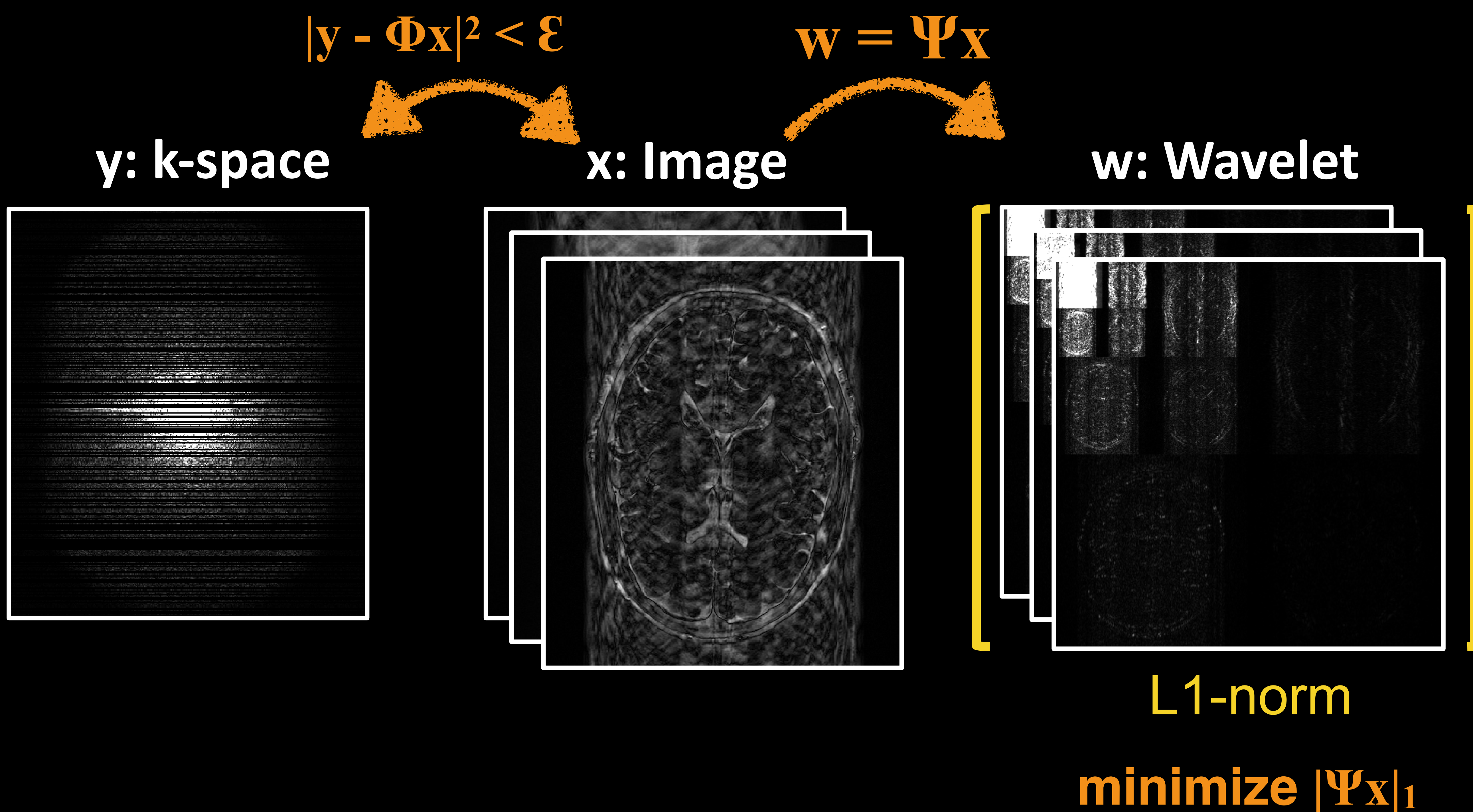
L0-norm ($|x|_0$): a number of non-zero coefficients

L1-norm ($|x|_1$): a sum of absolute values of coefficients

L2-norm ($|x|_2$): a sum of squared values of coefficients

$$
\begin{array}{ccc}
x & x & x \\
\begin{bmatrix} 0 \\ 1 \\ 2 \\ 3 \end{bmatrix} &
\begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} &
\begin{bmatrix} 1 \\ 1 \\ -2 \\ 3 \end{bmatrix}
\end{array}
$$

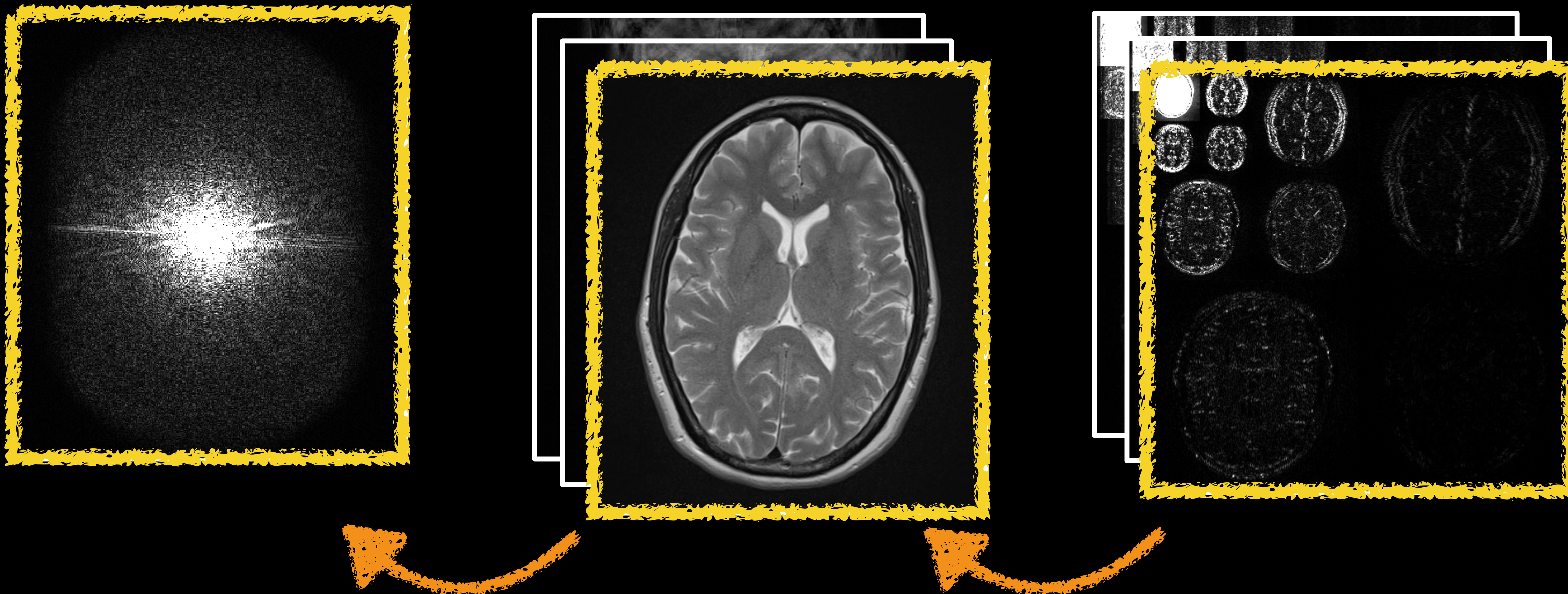
CS-MRI Reconstruction



CS-MRI Reconstruction

minimize $F(x)$: $|y - \Phi x|^2 + R(x)$

y: k-space **x: Image** **w: Wavelet**



$y' = FT(x)$ $x = \Psi^{-1} w$

Three Tenets of CS

minimize $F(x)$: $\underbrace{|y - \Phi x|_2^2}_{\text{Data Consistency}} + \underbrace{R(x)}_{\text{Compressibility Constraint}}$

- Three key elements of Compressed Sensing:

Compressibility

Incoherence

Nonlinear Reconstruction

Compressibility Constraint

minimize $F(x): |y - \Phi x|^2{}_2 + \underbrace{R(x)}_{}$

Compressibility Constraint

- $R(x) = \lambda|x|_1$ (*Identity Transform*)
- $R(x) = \lambda|\Psi x|_1$ (*Wavelet Transform*)
- $R(x) = \lambda H(x)$ (*Total Variation*)
- $R(x) = \lambda|x|_*$ (*Rank or Nuclear Norm*)
- Many more…

Wavelet Transform

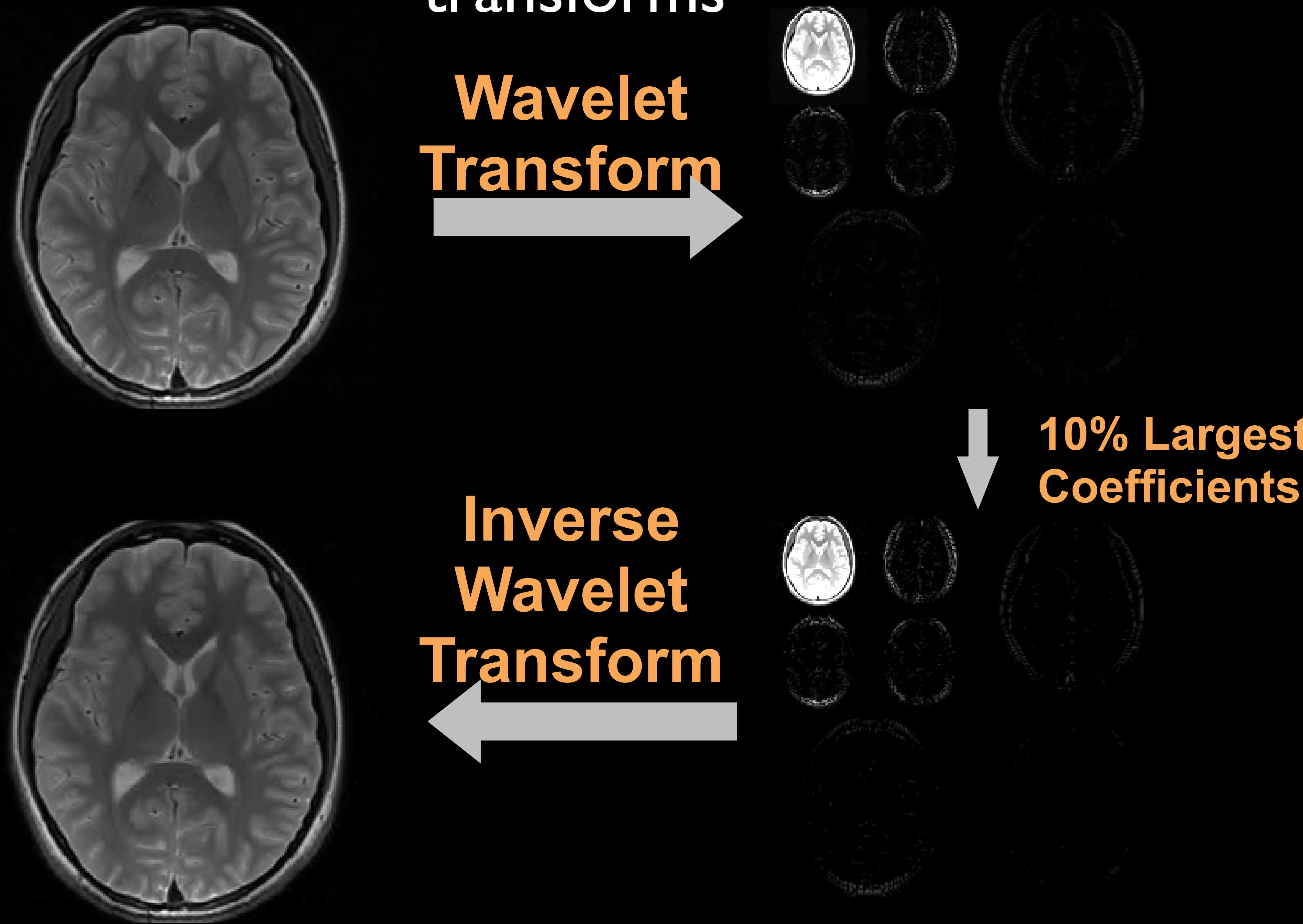
- Natural images are compressible using wavelet transforms

Image Compression Standard: JPEG2000

Images from Wikipedia

Wavelet Transform

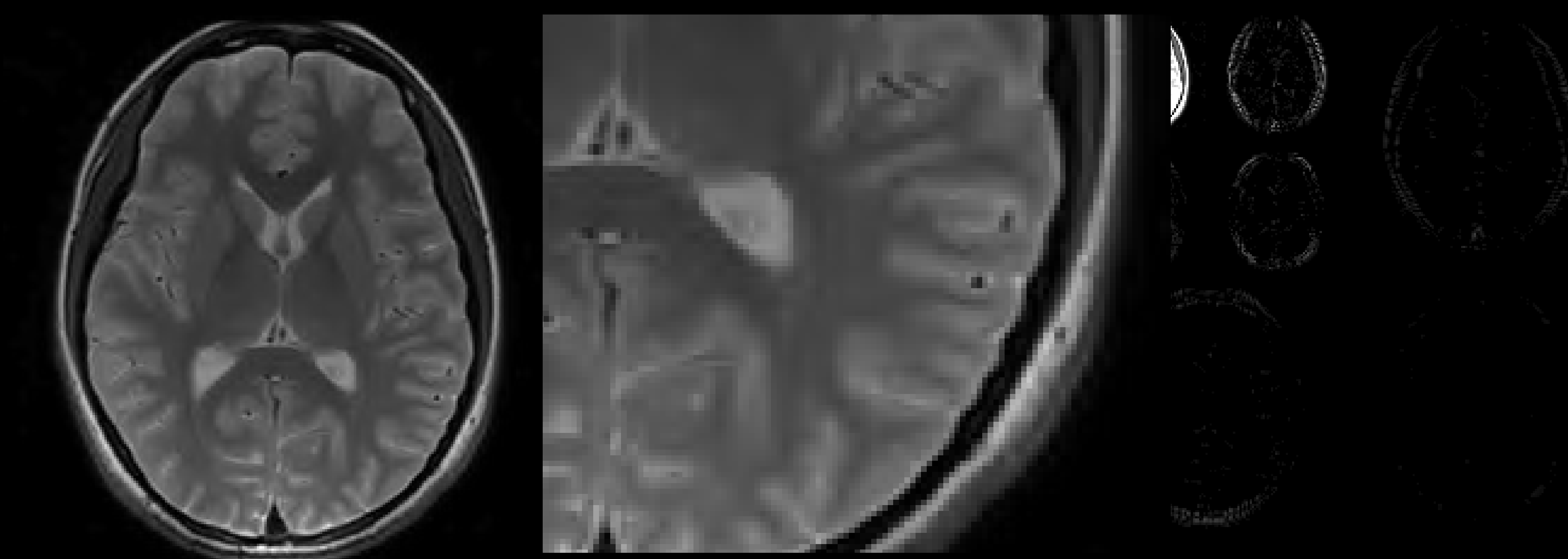
MR images are mostly compressible using wavelet transforms



Wavelet Transform

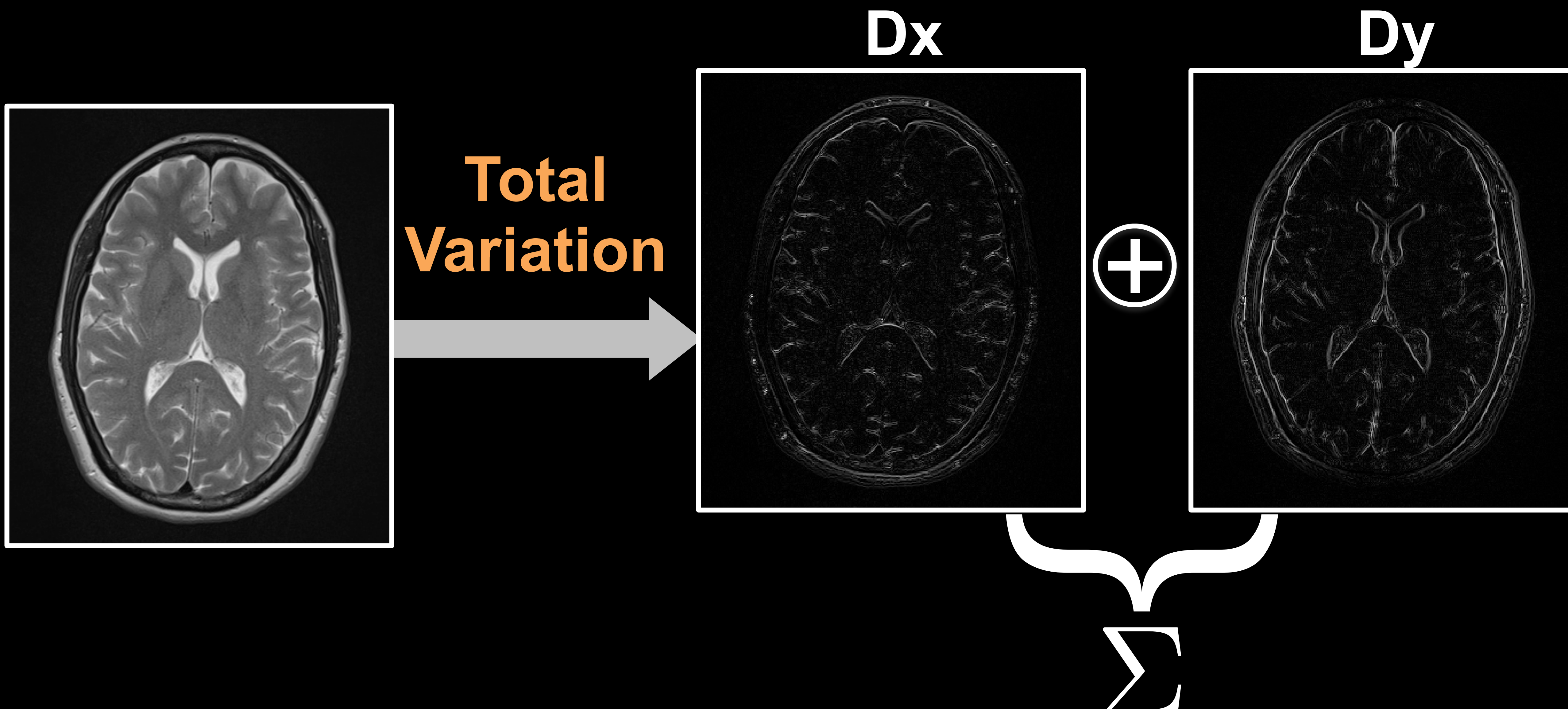
MR images are mostly compressible using wavelet transforms

10% Largest Coefficients

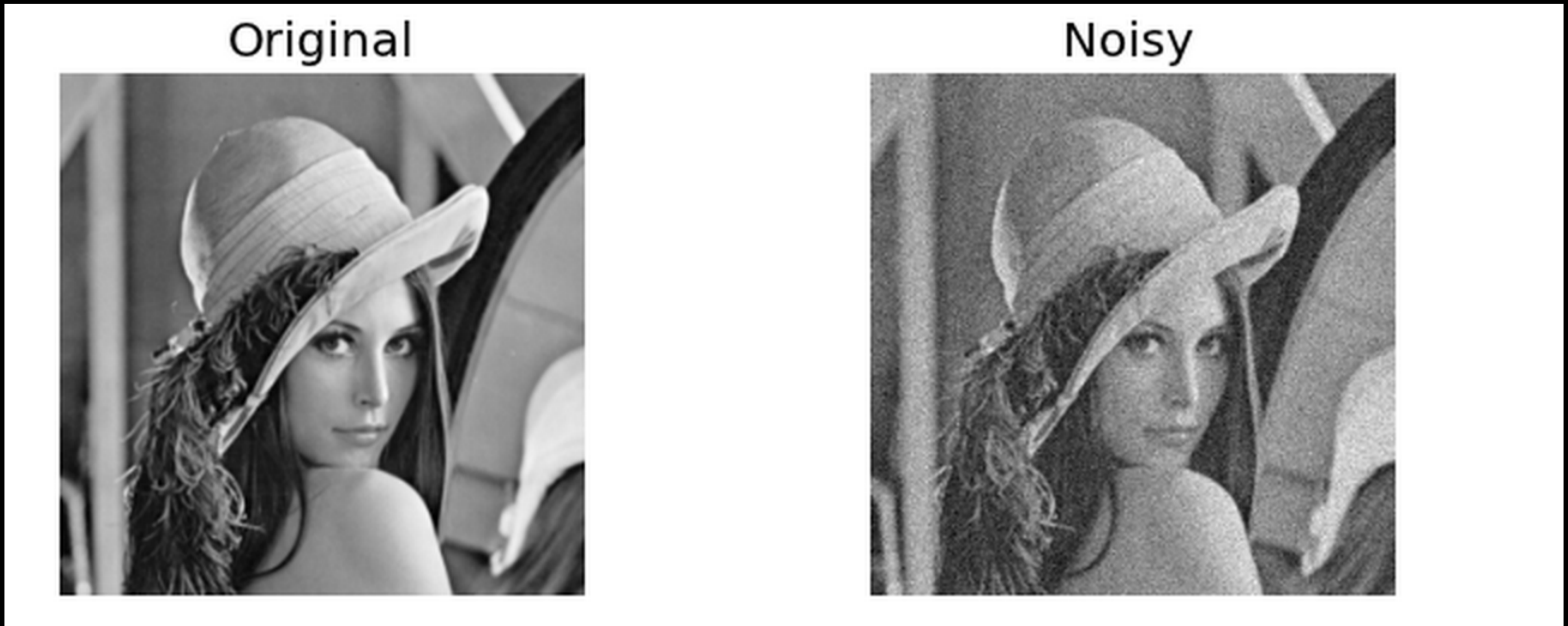


Total Variation

$$H(x) = \sum_{i,j} \sqrt{\underbrace{|x_{i+1,j} - x_{i,j}|^2}_{\mathbf{Dx}} + \underbrace{|x_{i,j+1} - x_{i,j}|^2}_{\mathbf{Dy}}}$$



Total Variation



Limitations / Considerations

- Define reconstruction domain and exploit information redundancy (or prior knowledge)
  - More apparent when MRI is repeated on a same object (e.g., repeating with different time points, flip angles, TEs, etc)
- Be aware of underlying assumptions of each constraint
  - Wavelet / TV denoising
- Consistent compressibility is desirable to easily anticipate reconstruction quality

CS-MRI Reconstruction

minimize $F(x): |y - \Phi x|_2^2 + R(x)$

- Minimizing F(x) is non-trivial since R(x) is not <u>differentiable</u>
  - Linear programming is challenging due to high computational complexity
- Simple gradient-based algorithms have been developed:
  - Re-weighted L1 / FOCUSS
  - IST / IHT / AMP / FISTA
  - Split Bregman / ADMM

I.F. Gorodnitsky, et al., J. Electroencephalog. Clinical Neurophysiol. 1995 Daubechies I, et al. Commun. Pure Appl. Math. 2004
Elad M, et al. in Proc. SPIE 2007
T. Goldstein, S. Osher, SIAM J. Imaging Sci. 2009

State-of-the-Art CS-MRI

- Reducing possible reconstruction failure
  - Improve sparse transformations
  - Develop k-space undersampling schemes
- Integrating CS with DL/parallel imaging
  - Develop compatible undersampling patterns
  - Develop reconstruction methods

State-of-the-Art CS-MRI

- Methods to evaluate CS reconstructed images
  - RMSE / SSIM / Mutual Information
- Reducing reconstruction time
  - Reduce computational complexity
  - Parallelize reconstruction problems
- Developing stable reconstruction algorithms
  - Minimize / avoid the number of regularization parameters

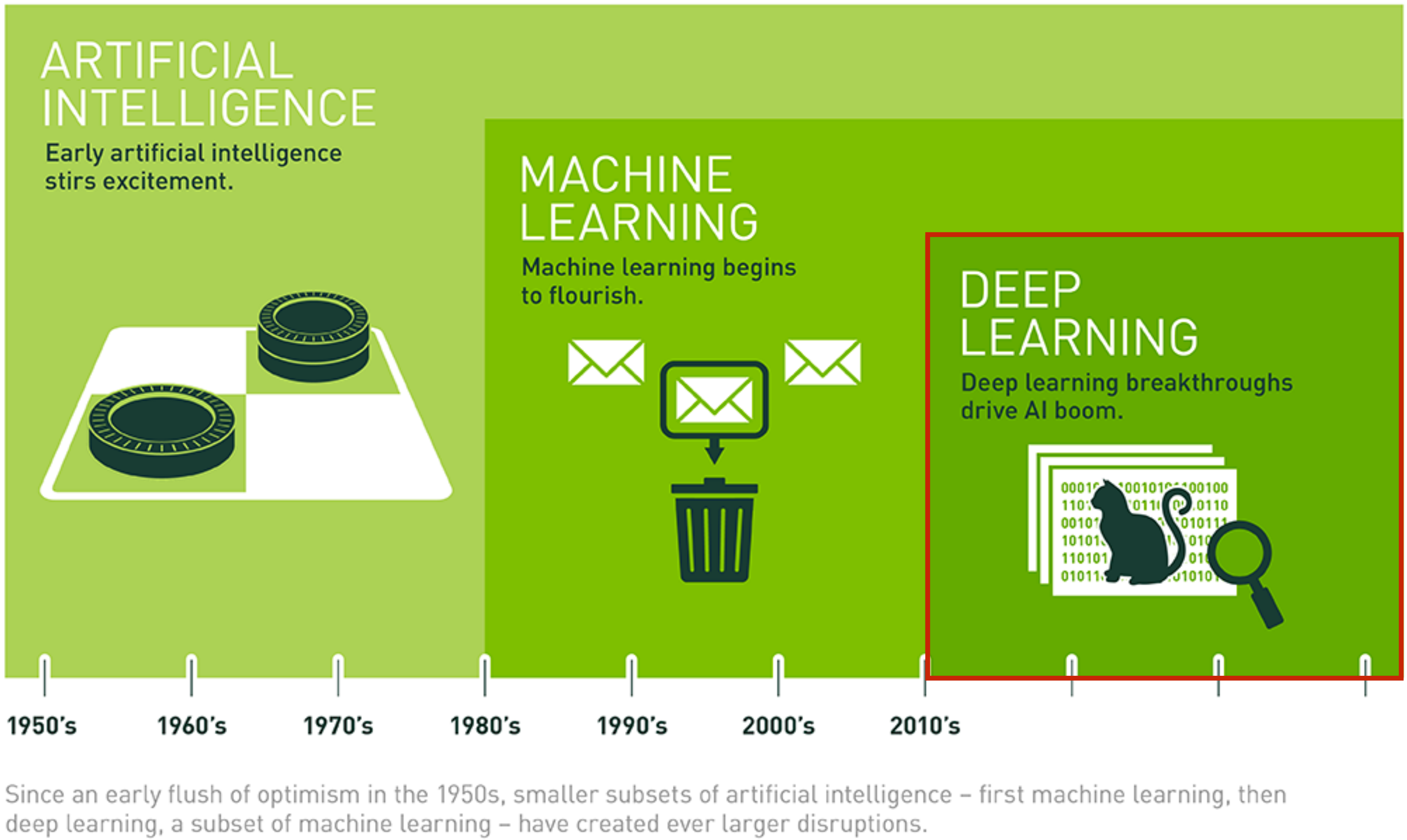
Summary So Far...

minimize $F(x)$: $\underbrace{|y - \Phi x|_2^2}_{\text{Data Consistency}} + \underbrace{R(x)}_{\text{Compressibility Constraint}}$

Compressibility Constraint

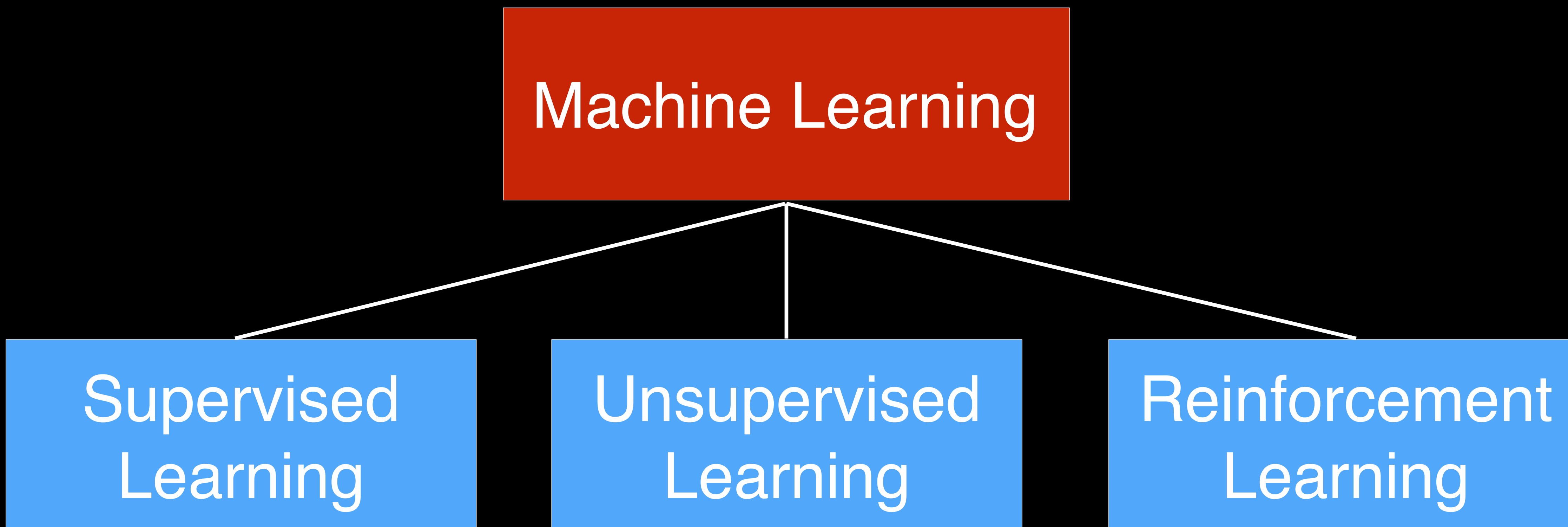
Incoherent Measurement

Reconstruction



From nvidia.com

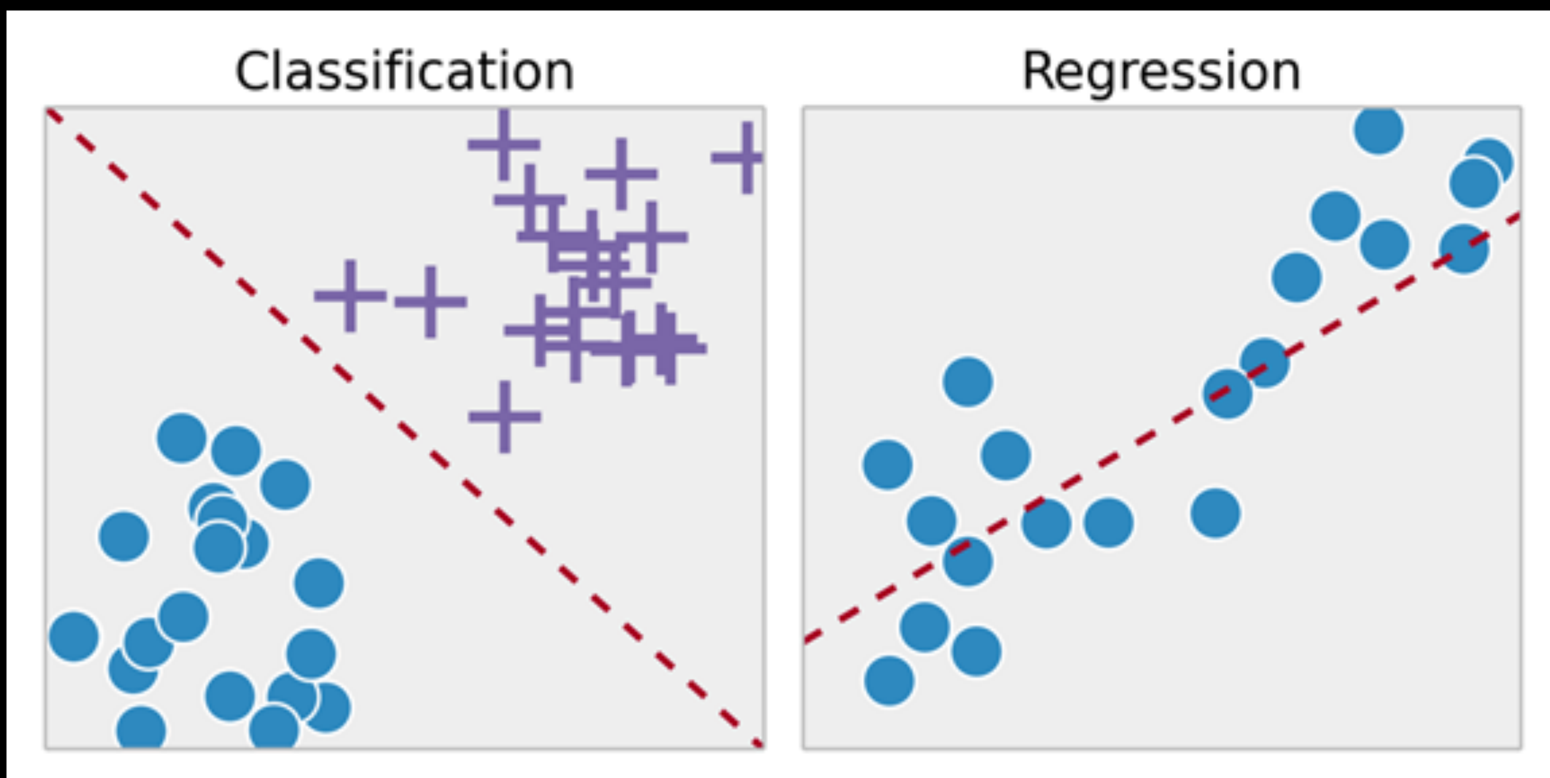
Types of Machine Learning



Supervised Learning	Unsupervised Learning	Reinforcement Learning
Task driven (Regression /Classification)	Data driven (Clustering)	Algorithm learns to react to an environment

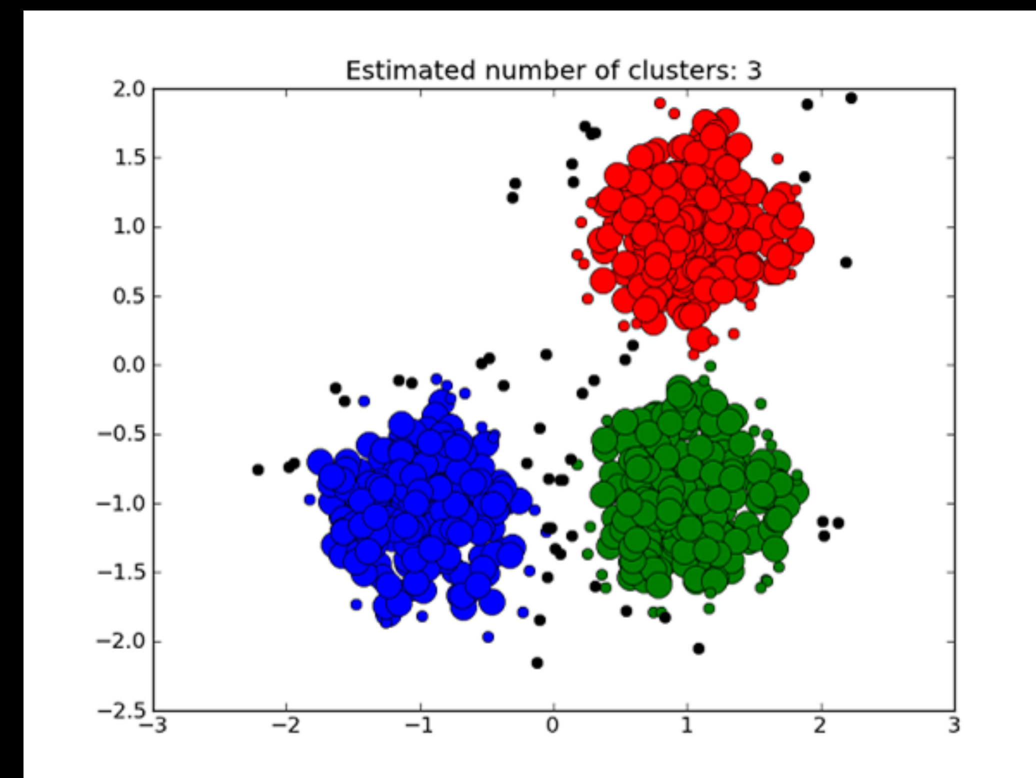
Types of Machine Learning

Supervised Learning



Discriminative
Model

Unsupervised Learning



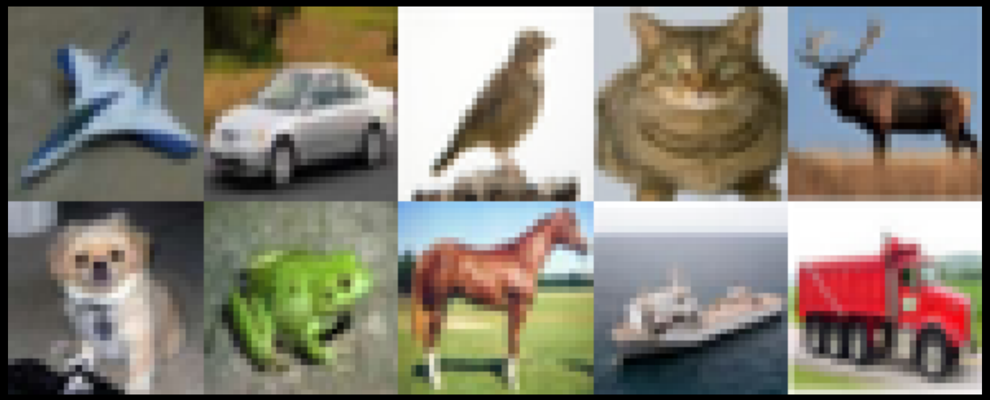
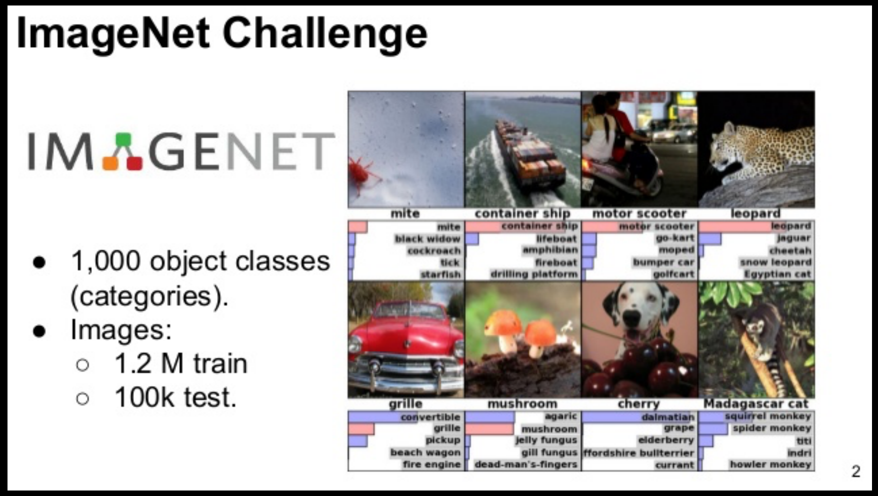
Generative Model

Common Datasets for Deep Learning

- MNIST: 60,000 images, hand writing digits (1998)
- CIFAR-10: 60,000 images 10 classes of

Not trivial to build medical imaging database with a high number of images and accurate labeling

- ImageNet: 1,300,000 high-res images, 1,000 classes of object (2012)

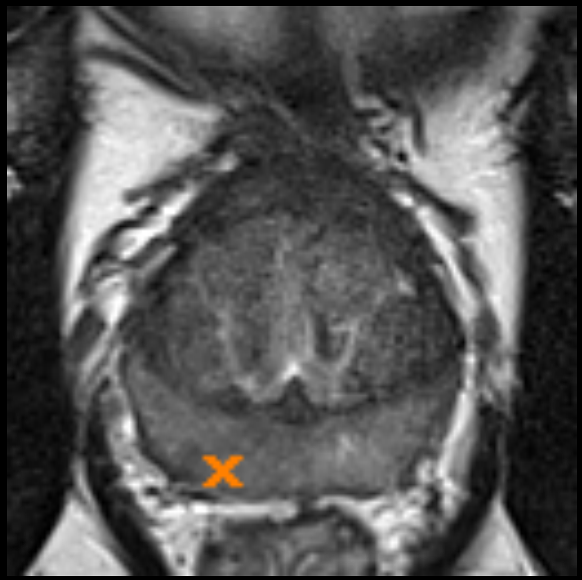




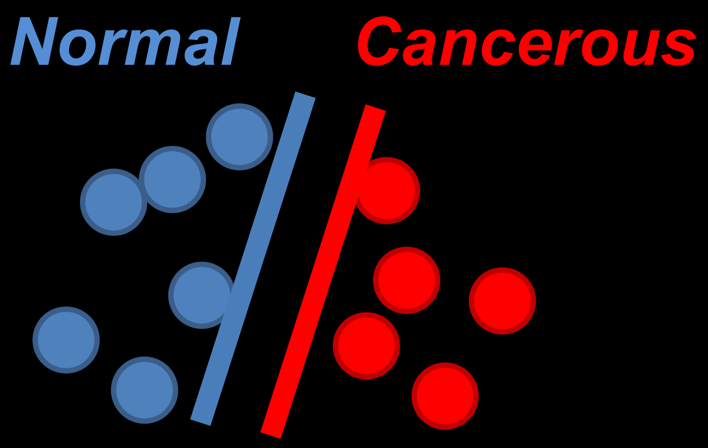
Key Design Considerations

1. Define clear clinical questions
2. Design deep learning models
   - Supervised vs. unsupervised learning
   - Descriptive vs. generative modeling
3. Consider potential limitations
   - Limited amount of training and testing data
   - Uncertainties in labeling

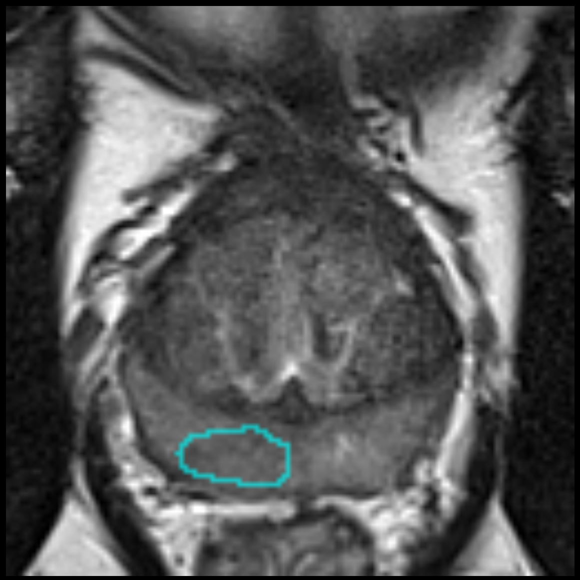
Artificial Intelligence for MRI

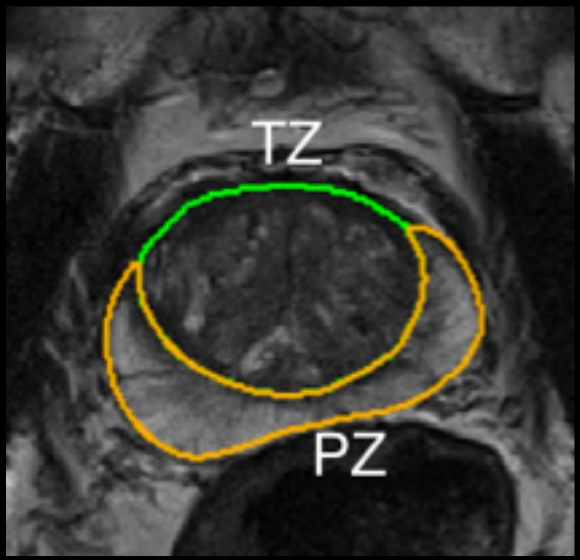
Detection



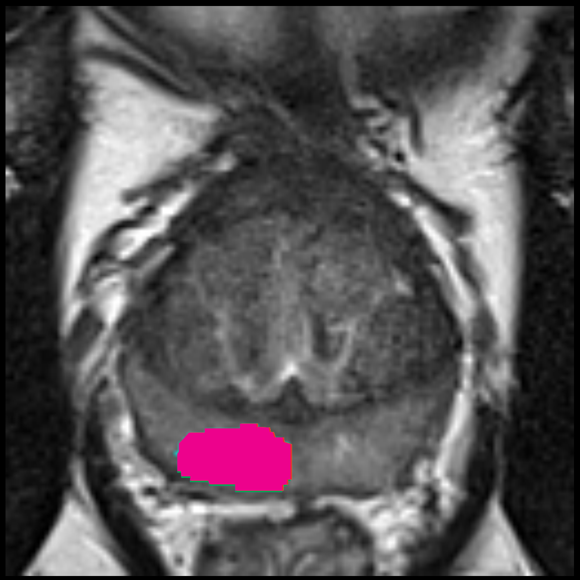


Segmentation





Classification



Histologic Findings

GS 3+3

GS 3+4

GS 4+3

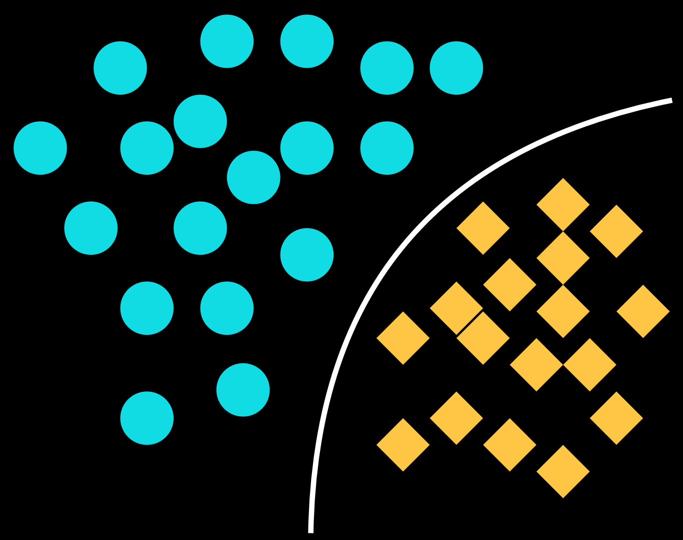
GS=8

GS=9 or above

Model Evaluation, Model Selection, and Algorithm Selection

- <u>Target function</u> is a specific, unknown model that we want to learn or approximate
- <u>Model</u> is a certain function that we believe is similar to the true function, the target function that we want to model
- <u>Learning algorithm</u> is a set of instructions that tried to model the target function using a training dataset
- <u>Hyperparameters</u> are the tuning parameters of a machine learning algorithm

Target Function and Model



Target function
$y = f(x)$

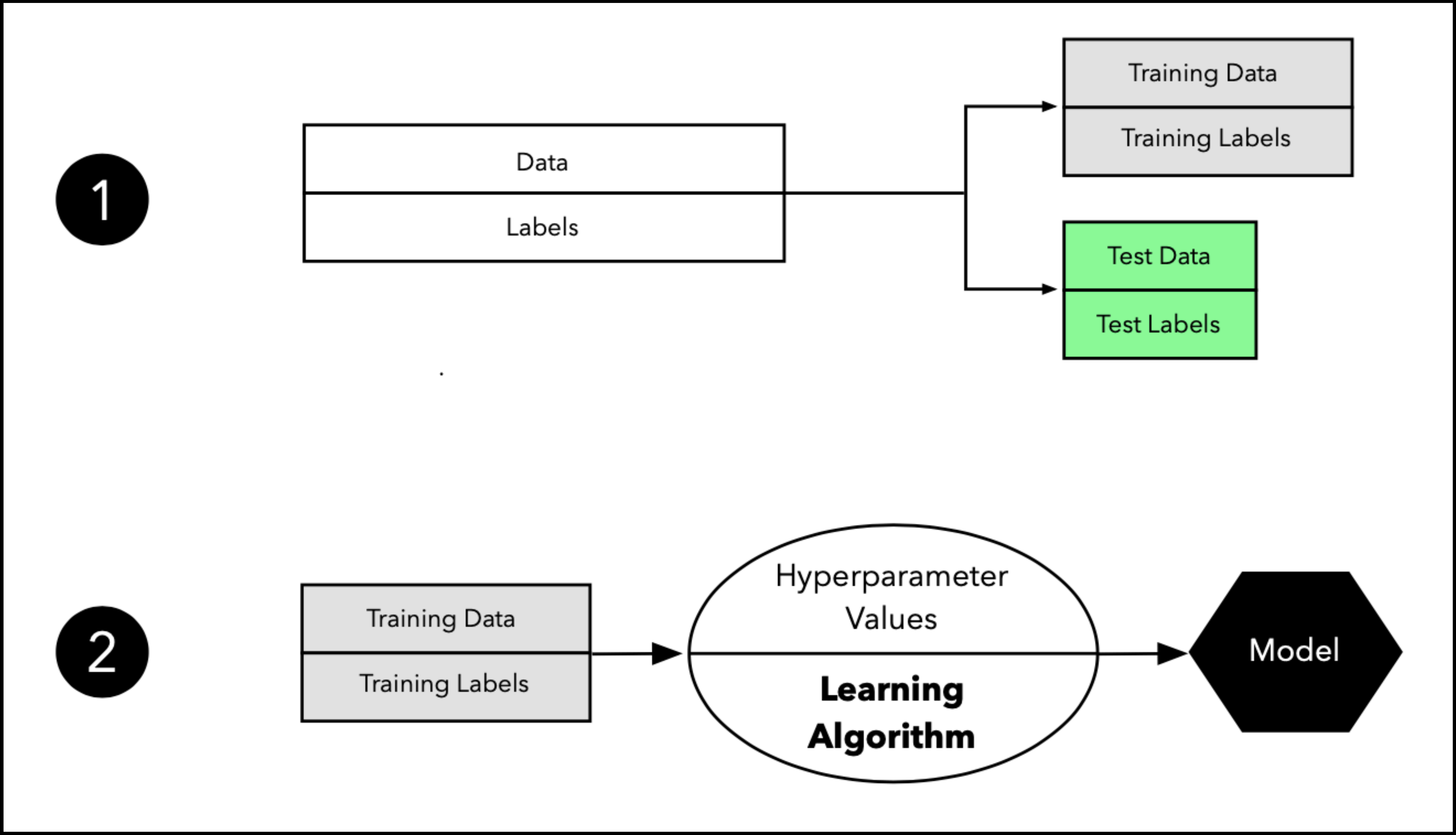
Model: $y = ax + b$

Model parameters: a and b

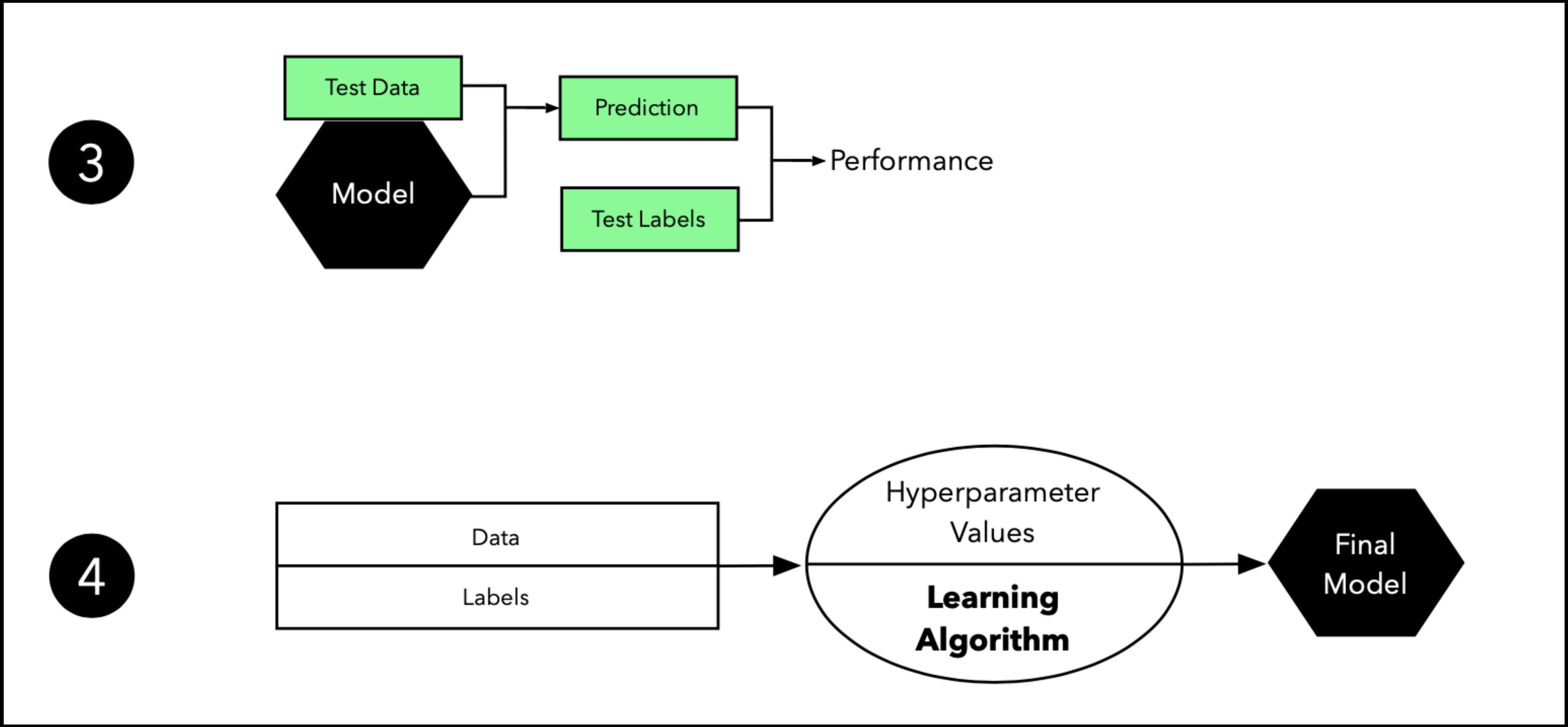
Learning Algorithm

Hyperparameters

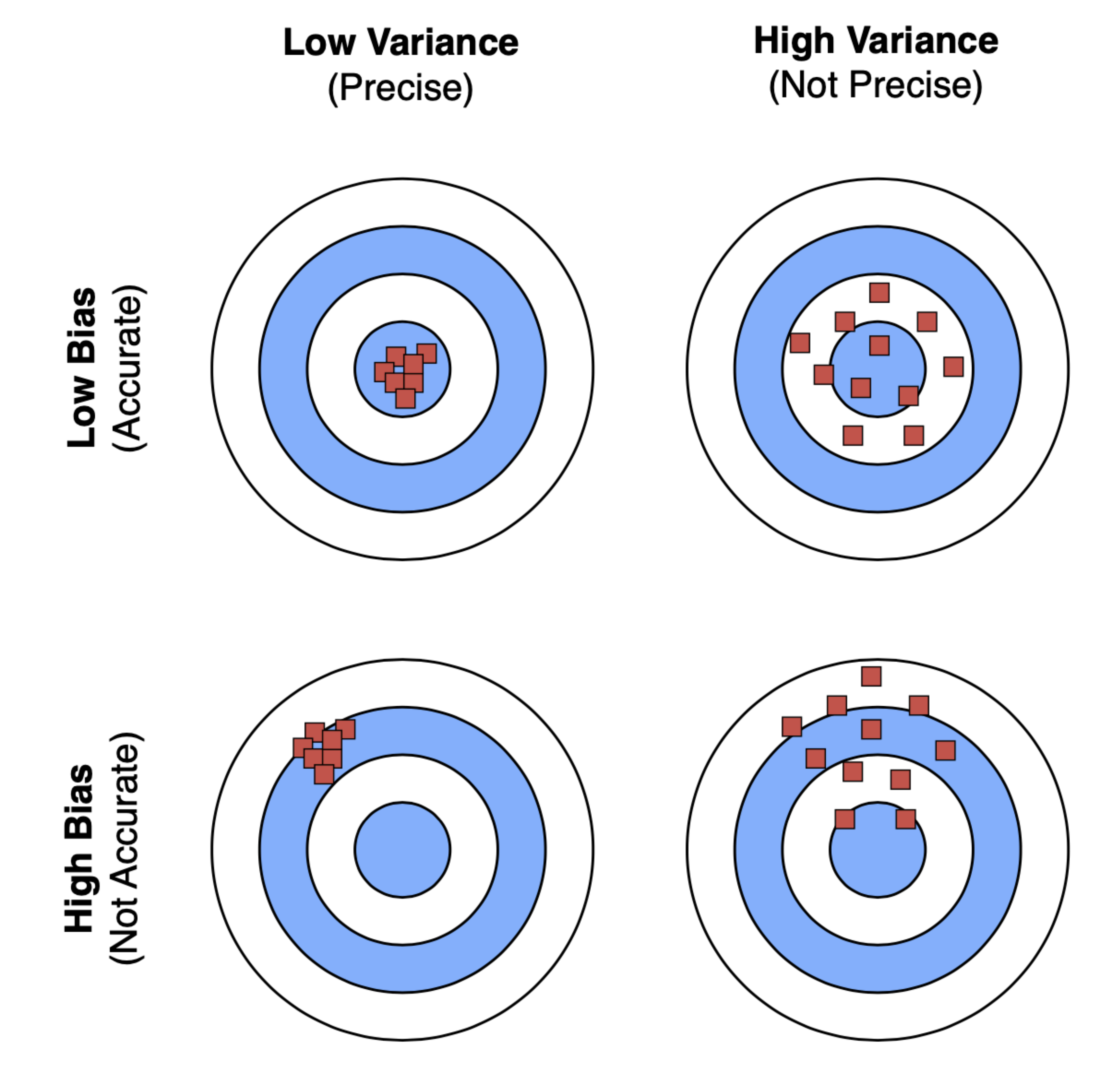
Holdout Validation Method



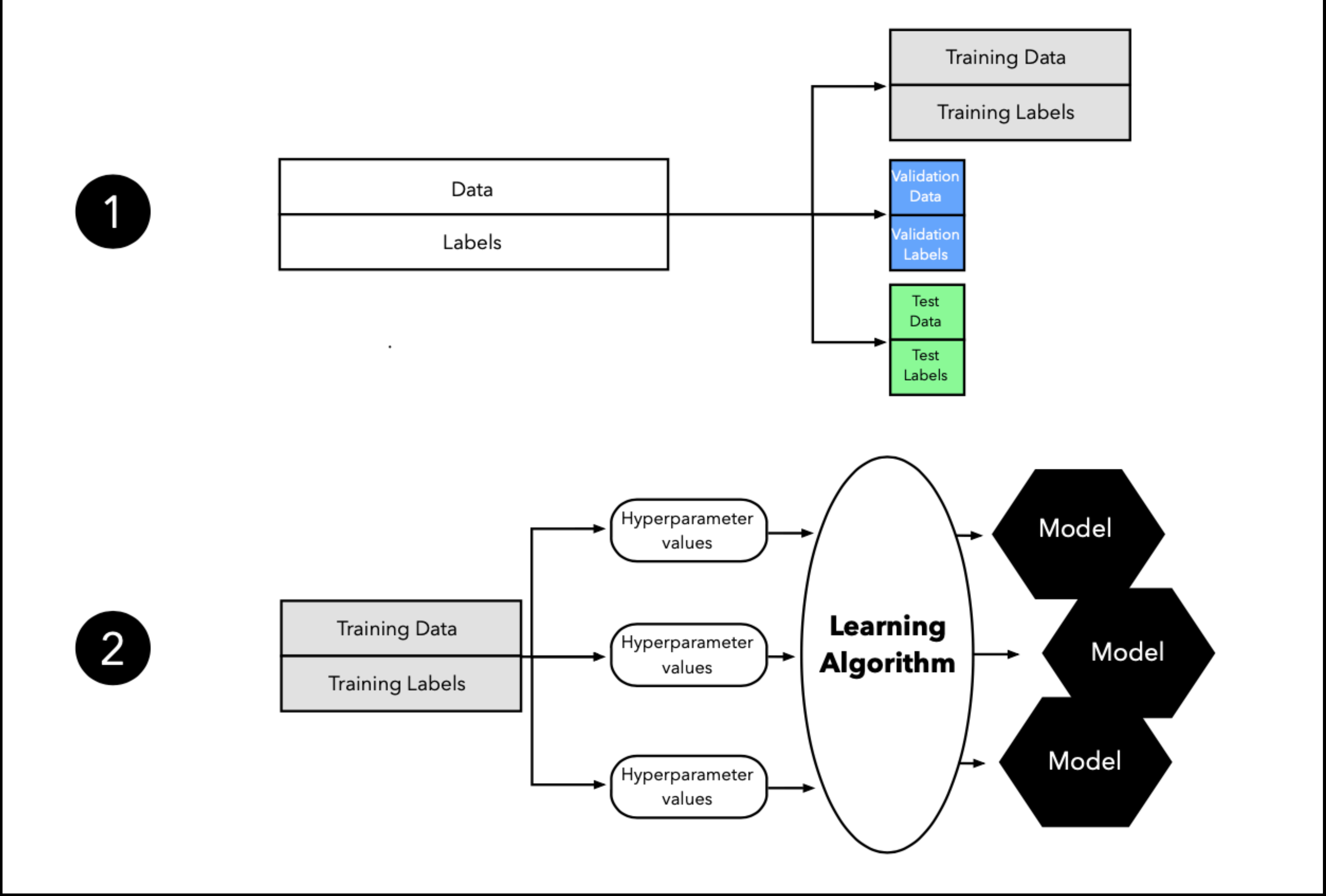
Holdout Validation



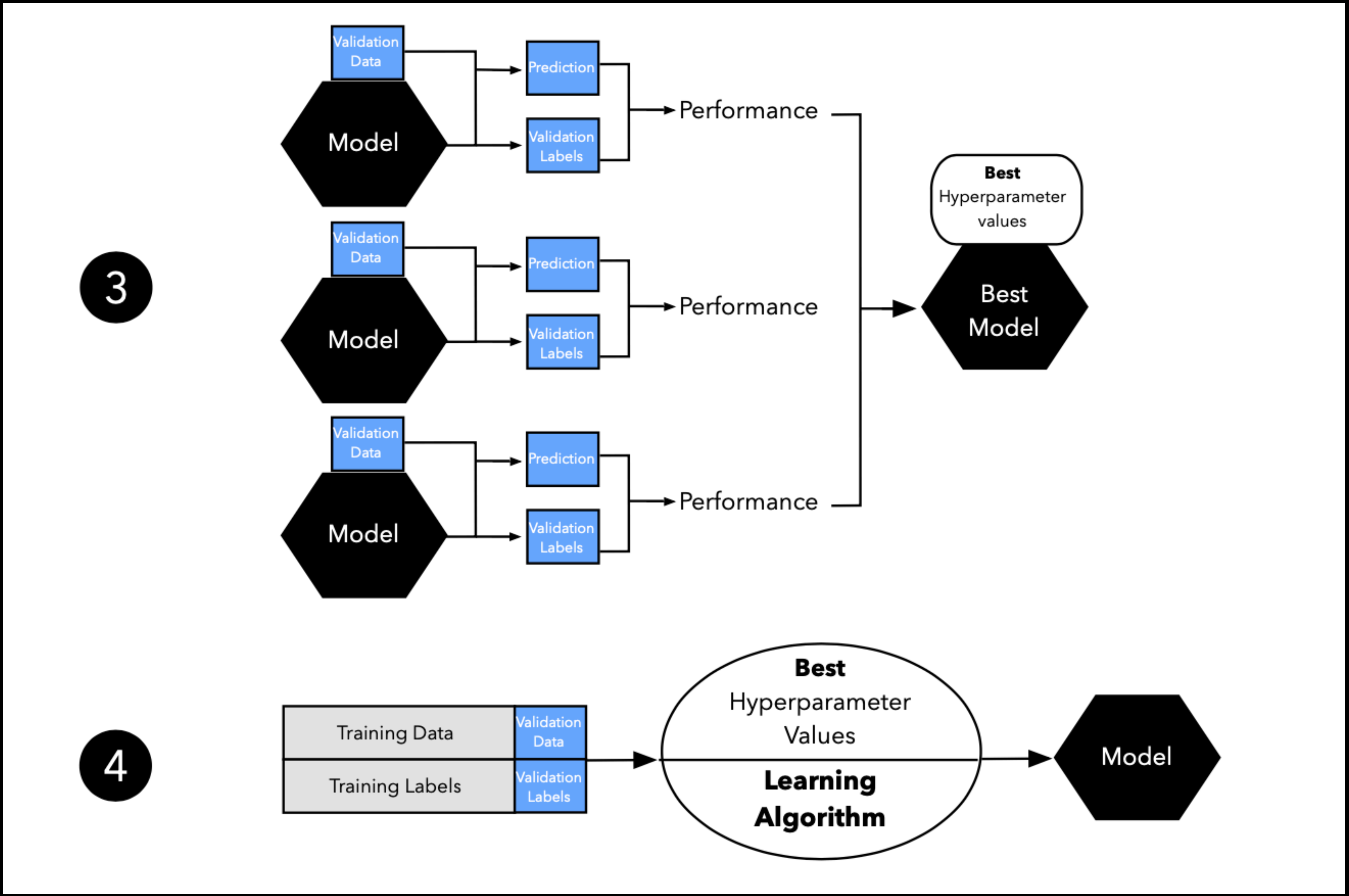
Bias and Variance



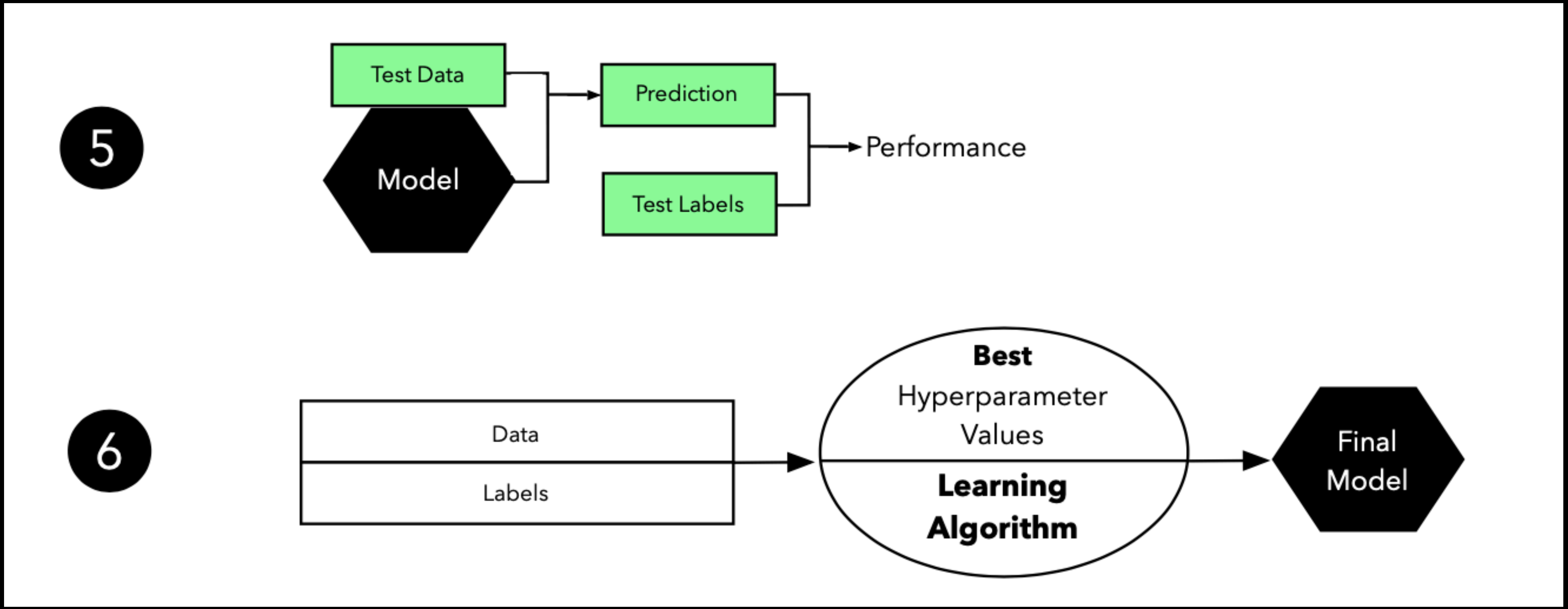
Three-way Holdout Method



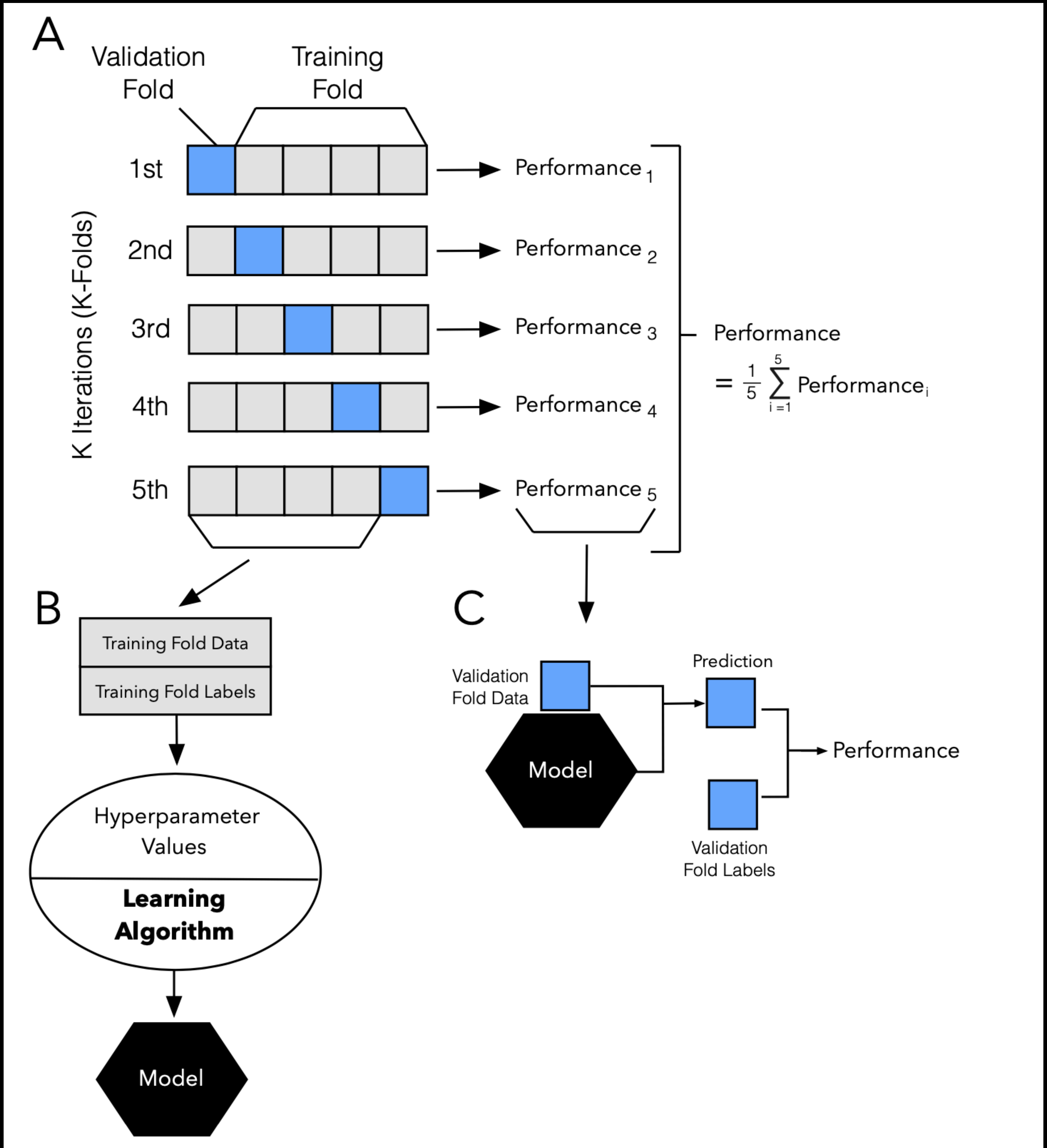
Three-way Holdout Method



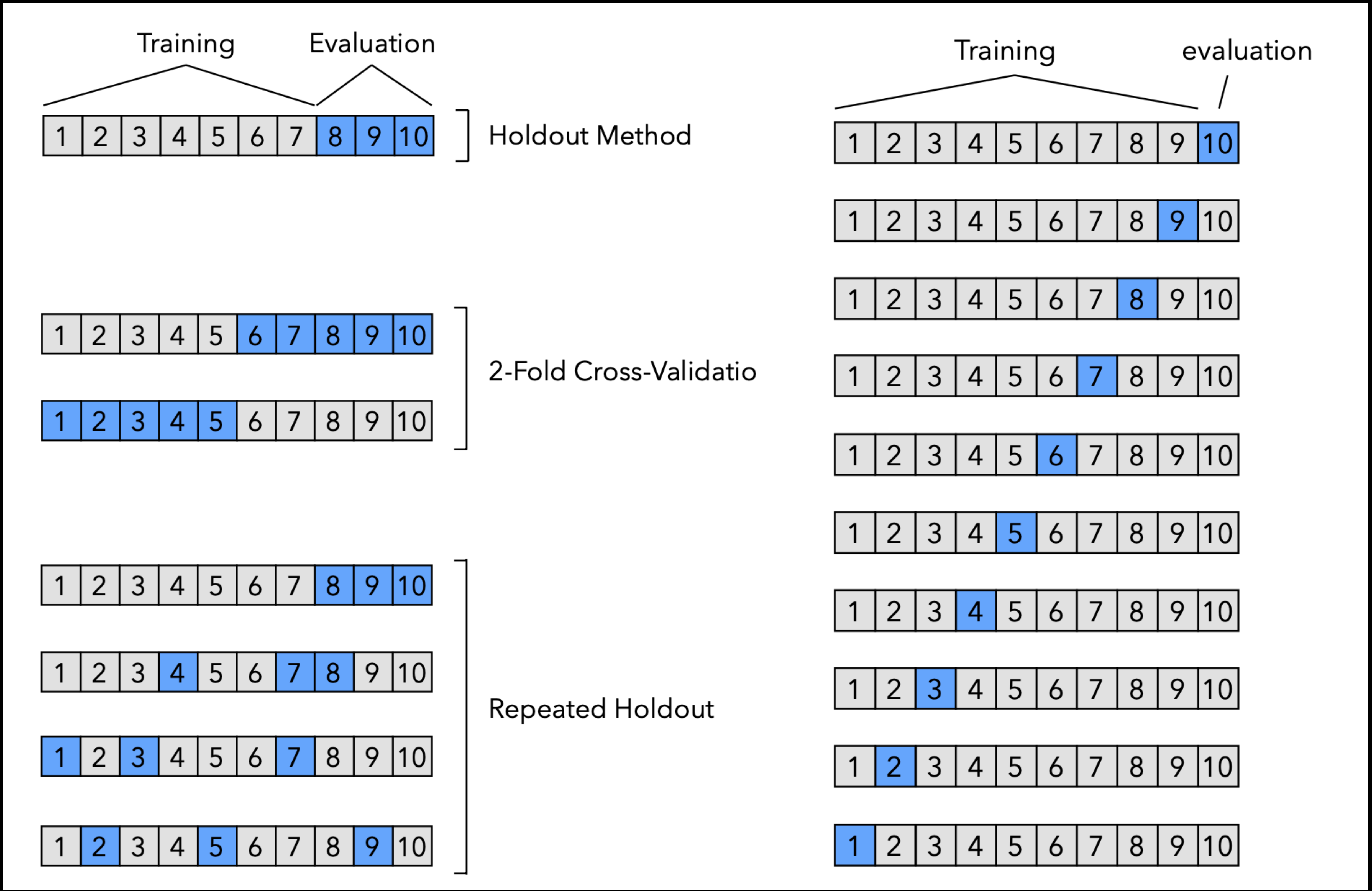
Three-way Holdout Method



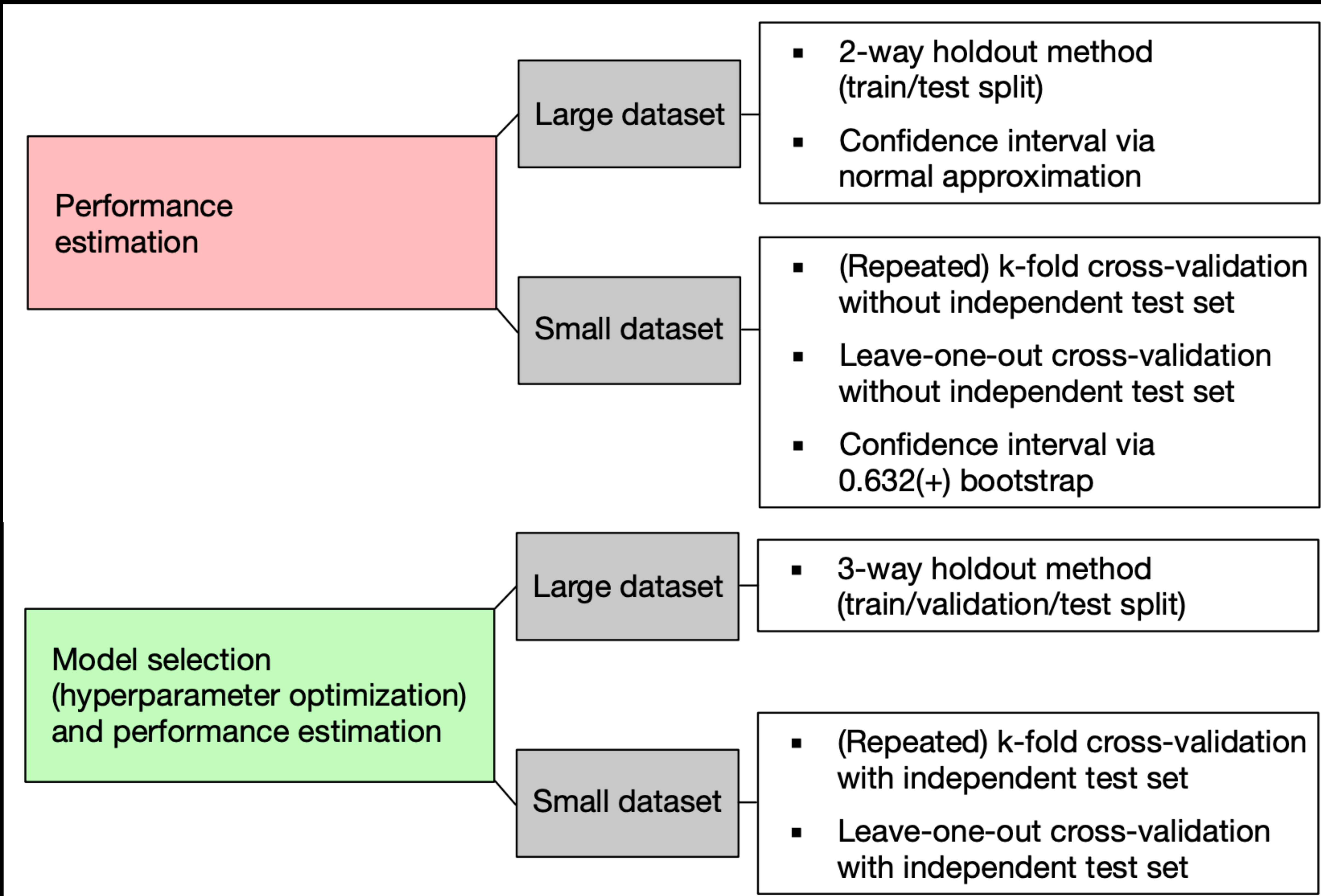
k-fold Cross-Validation



Holdout / CV / Repeated Holdout / LOOCV



Summary



Performance estimation
Large dataset
2-way holdout method (train/test split)
Confidence interval via normal approximation
Small dataset
(Repeated) k-fold cross-validation without independent test set
Leave-one-out cross-validation without independent test set
Confidence interval via 0.632(+) bootstrap
Model selection (hyperparameter optimization) and performance estimation
Large dataset
3-way holdout method (train/validation/test split)
Small dataset
(Repeated) k-fold cross-validation with independent test set
Leave-one-out cross-validation with independent test set

Further Reading

- Original Compressed Sensing
  - https://ieeexplore.ieee.org/document/1580791
  - https://ieeexplore.ieee.org/document/1614066
- Compressed Sensing MRI
  - https://ieeexplore.ieee.org/abstract/document/4472246
- ML model selection and evaluation
  - https://arxiv.org/abs/1811.12808

Thanks!

- Next time
  - Dr. Yingli Yang (6/2)
  - Dr. Fabien Scalzo (6/4)

Kyung Sung, PhD

ksung@mednet.ucla.edu

https://mrrl.ucla.edu/sunglab/