
University of Colorado
Boulder

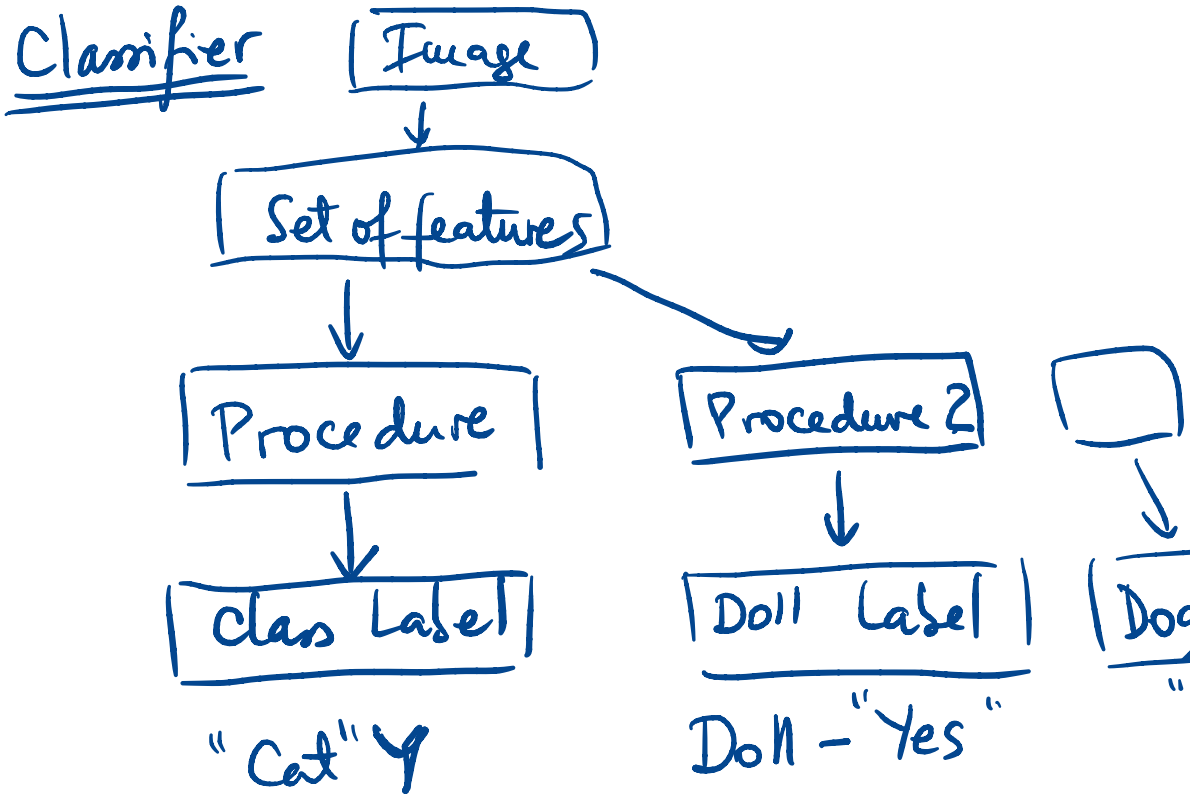
Deep Learning Applications for Computer Vision

Lecture 11: Object Recognition and Image Classification with Neural Networks

Object Recognition vs Image Classification

- Often used interchangeably

IC = label(s)



cat, doll, book, bed

Classification

raw image / feature vector

label

(x_i, y_i)

We need:

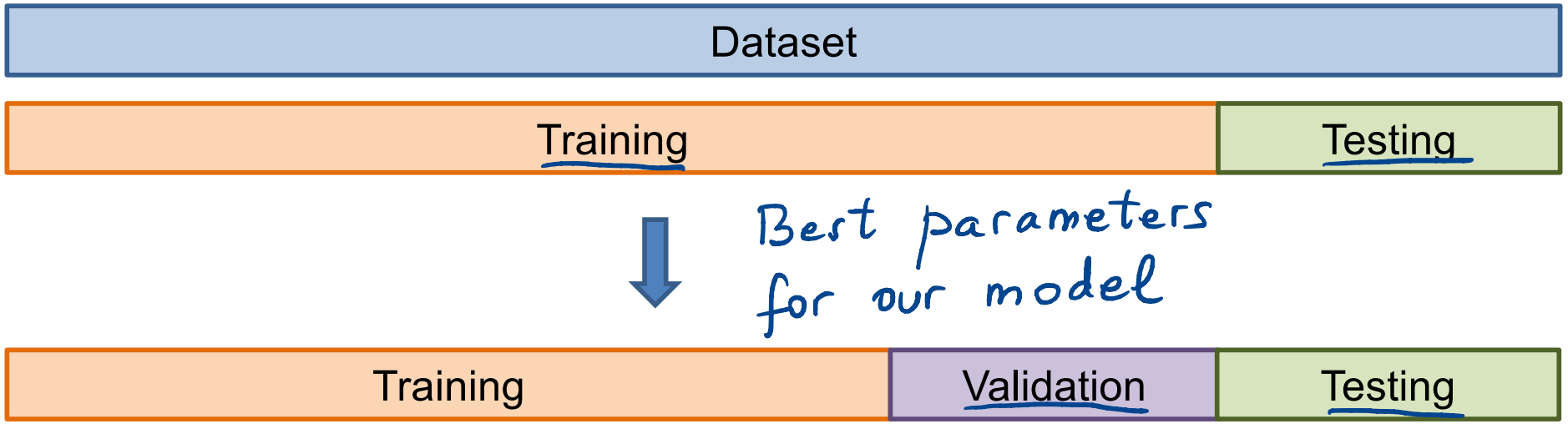
- Data = set of labeled examples (images)
- Rule/procedure: generates label for new image

Prediction function

$f($  $) =$ "cat"

$f(x) = y$

The Data



Min. prediction error → assigns the wrong label

Ex: 1000 images in Training dataset labeled "Dog"

950 predicted "Dog"

Training error = 5%

Confusion Matrix

Assumption: two-class classifier < Dog / Cat / Not Dog

- OL = original label
- PL = predicted label

		Predicted Condition → Model Predicts	
		Positive (PP) "Dog"	Negative (PN) ("Not Dog")
Actual Condition	Positive (P) "Dog"	True Positive (TP) "Dog" labeled as "Dog"	False Negative (FN) "Dog" → "Not dog"
	Negative (N) "Not Dog"	False Positive (FP) "Not Dog" → "Dog"	True Negative (TN) "Not Dog" → "Not dog"

$\text{Sensitivity} = \frac{TP}{P}$

$\text{Specificity} = \frac{TN}{N}$

Cost → Loss Function

What if the image is labeled incorrectly?

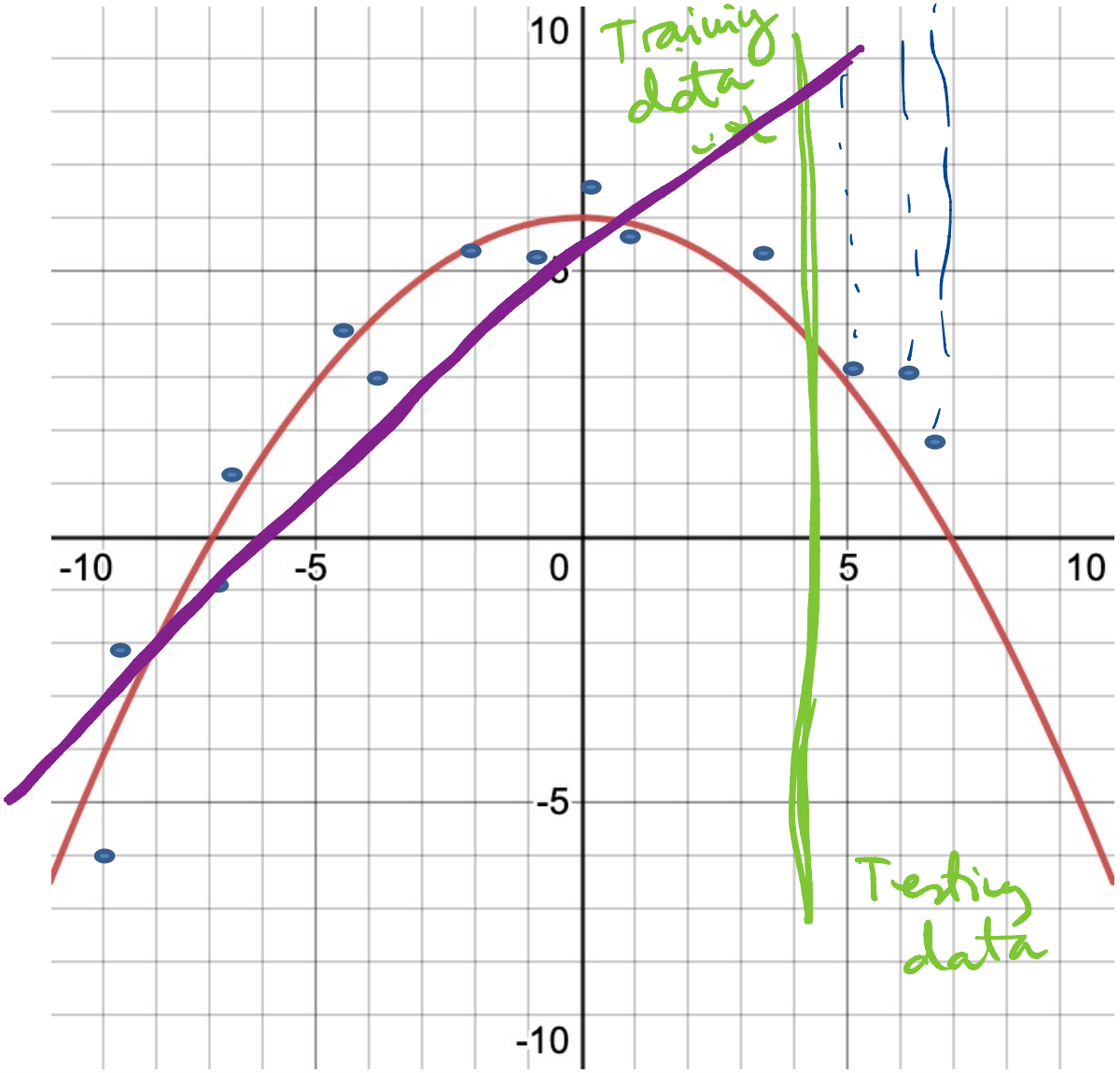
- **Loss function**: *L (i → j)* object of type **i** classified as type **j**
  - L(dog → dog) = 0
  - L (cat → dog) = … ?

- Risk function, **Total risk** – for strategy *s*

R(s) = p(dog → cat | using s)L(dog → cat) + FN

p(cat → dog | using s)L(cat → dog) FP

Generalization: Overfitting, Underfitting



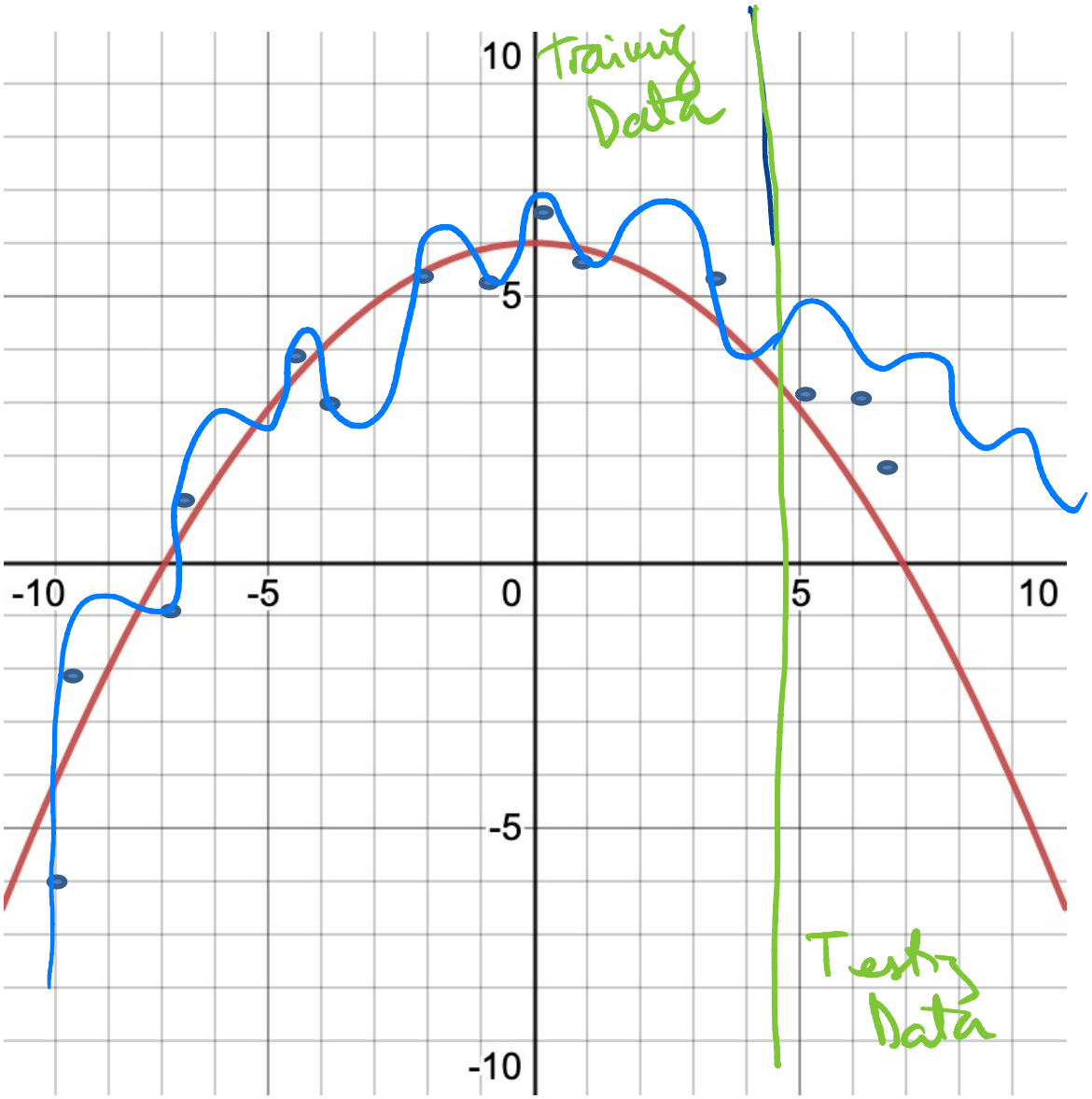
- **Underfitting**:
  - too few parameters
  - not enough flexibility

$f(x) = 6 - \frac{x^2}{8}$

Red → perfect fit

Blue dot → data + noise

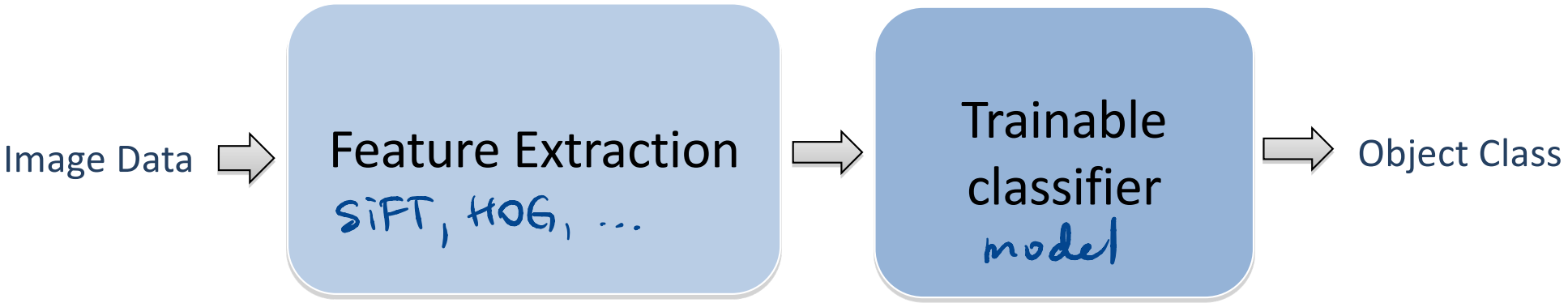
Generalization: Overfitting, Underfitting



- **Overfitting**:
  - too many parameters
  - trying too hard to have 0 loss for every testing sample

Classification – Traditional Approach

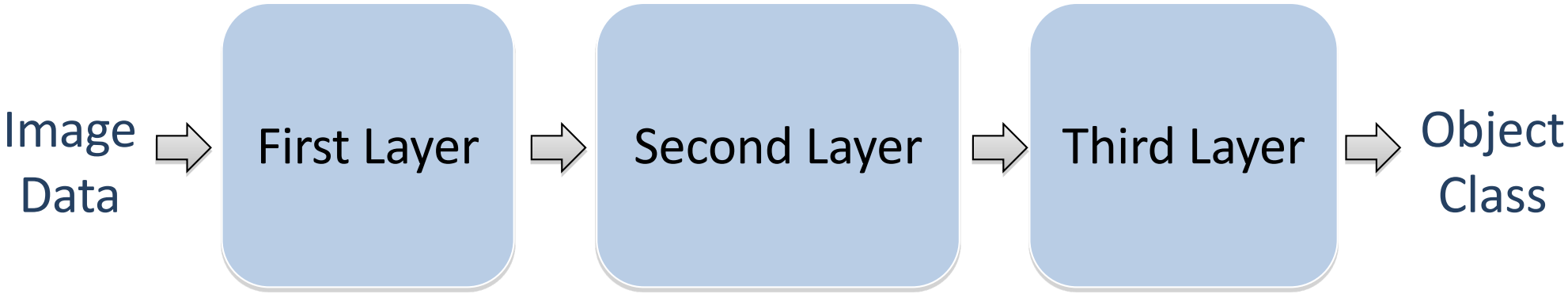
- relies on hand-designed feature extraction



Classification – Can we learn the features?

- Each layer learns from the output of the previous one
- Training is done for all layers together

Layer hierarchy



- How many layers???

Shallow vs. deep ⟩ architecture

Summary

- Image classification two-class vs multi-class
- Data and the cost of misclassification
- Shallow vs Deep architecture