LECTURE 12

SPRING 2021

APPLIED MACHINE LEARNING

CIHANG XIE

SLIDE CREDIT:
NARGES NOROUZI

QUIZ

- Bi-weekly, usually release on Tuesday
- Due on the same day
- FIVE in total (30%)

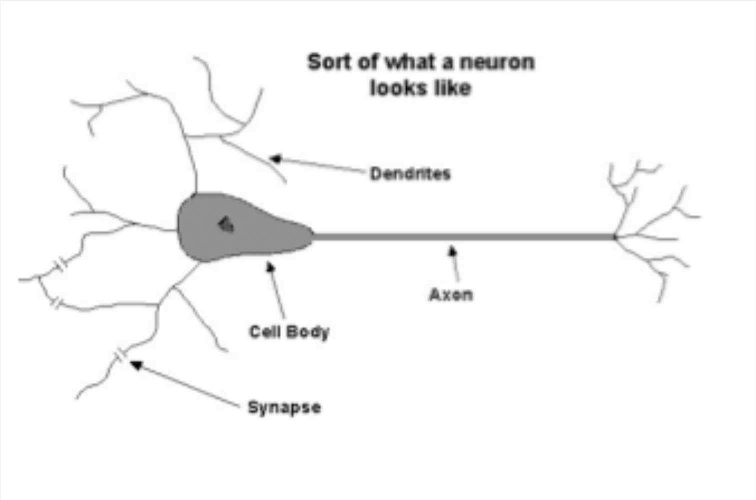
TODAY

- Artificial Neurons
  - Different activation functions
  - Implementation of AND, OR, XOR gates (functions)
- Introduction to Neural Networks

ARTIFICIAL NEURON

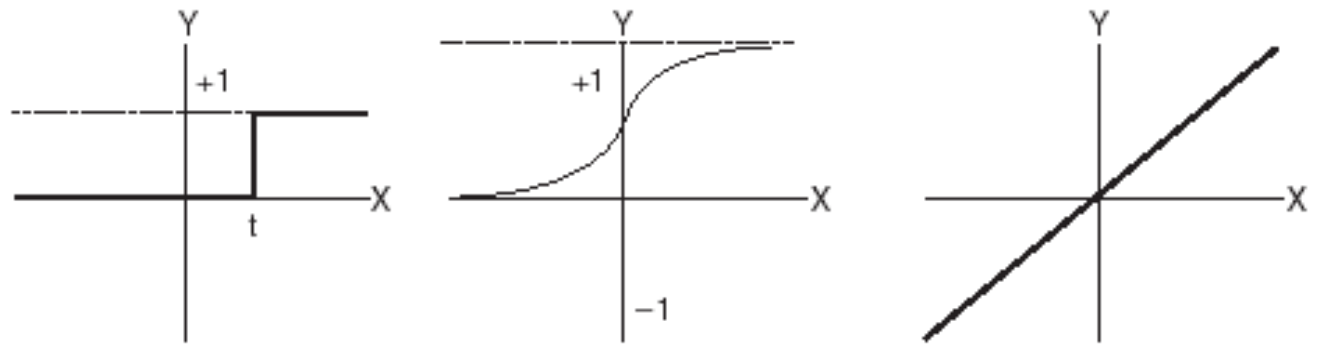
THE NEURON METAPHOR

- Neurons
  - Accept information from multiple inputs,
  - Transmit information to other neurons.
- Multiply inputs by weights/parameters along edges
- Apply some function to the set of inputs at each node



ARTIFICIAL NEURONS

- Each neuron receives one or more inputs.
- An activation function is applied to the inputs (net incoming activation), which determines the output of the neuron – the activation level.



(a) Step funtion (b) Sigmoid funtion (c) Linear funtion

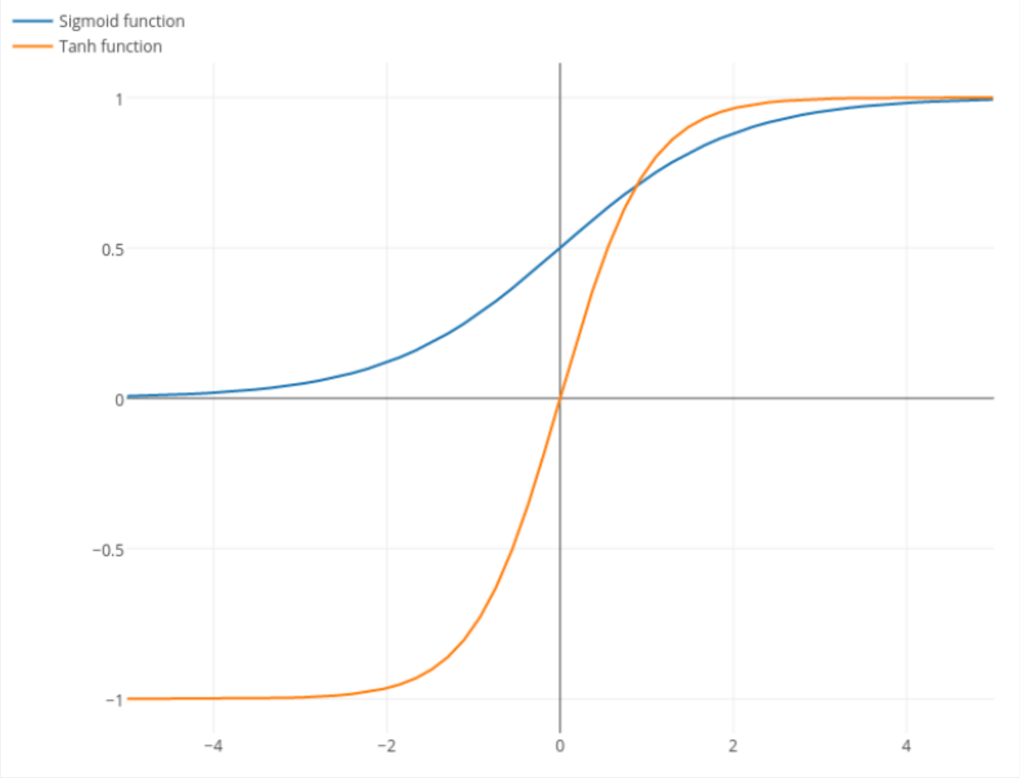
STANDARD ACTIVATION FUNCTIONS

- The hard-limiting threshold function
  - Corresponds to the biological paradigm
    - either fires or not

$$y = f(z) = \begin{cases} 0 & \text{if } z < t \\ 1 & \text{if } z \geq t \end{cases}$$

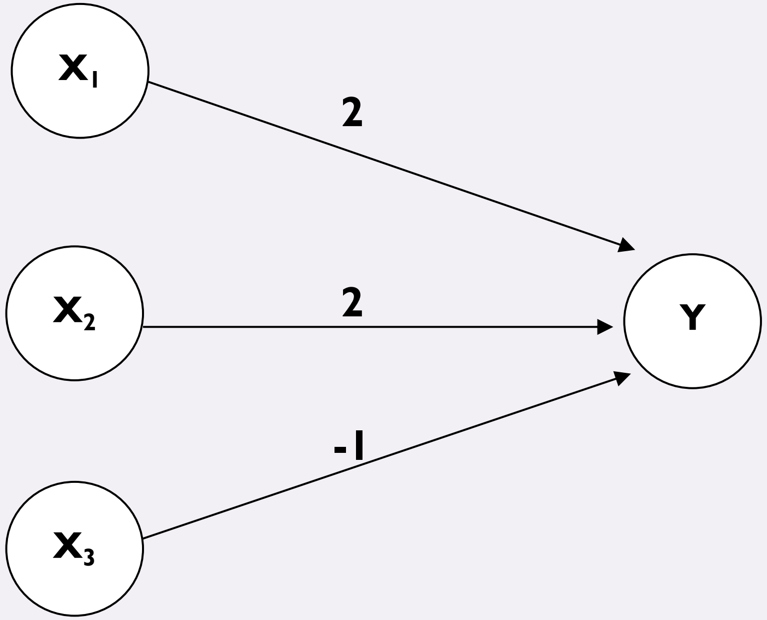
- Sigmoid function or the logistic function

- Hyperbolic tangent



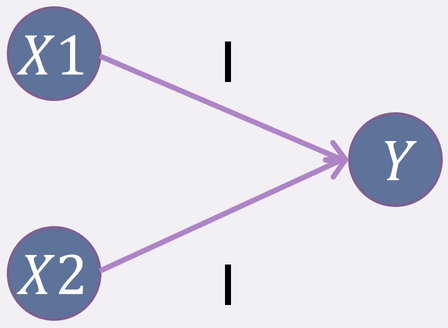
EXAMPLE:
STEP FUNCTION ACTIVATION

STEP FUNCTION PROPERTIES



- ✓ Assume that the activation of neuron is binary (not true in general). That is, the neuron either fires (activation of one) or does not fire (activation of zero).
- ✓ If the weight on a path is positive the path is excitatory, otherwise it is inhibitory.
- ✓ Each neuron has a fixed threshold. If the net input into the neuron is greater than the threshold, the neuron fires. The threshold is set such that any non-zero inhibitory input will prevent the neuron from firing.

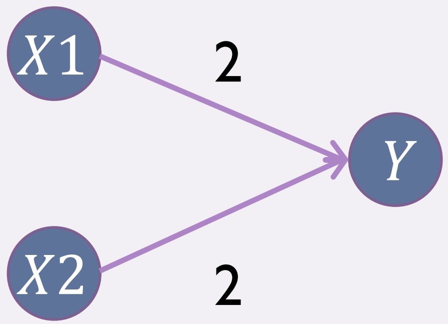
EXAMPLE - AND



AND Gate using step function with
$Threshold\ (Y) = 2$

AND		
X1	**X2**	**Y**
1	1	1
1	0	0
0	1	0
0	0	0

EXAMPLE - OR



OR Gate using step function with Threshold $(Y) = 2$

OR		
X1	**X2**	**Y**
1	1	1
1	0	1
0	1	1
0	0	0

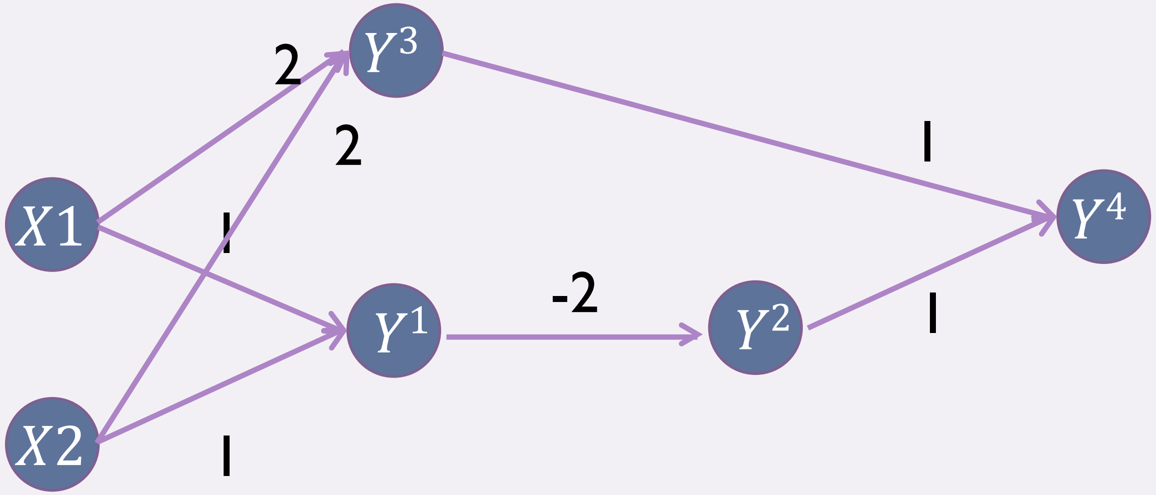
EXAMPLE - XOR

AND		
X1	X2	Y
1	1	1
1	0	0
0	1	0
0	0	0

OR		
X1	X2	Y
1	1	1
1	0	1
0	1	1
0	0	0

XOR		
X1	X2	Y
1	1	0
1	0	1
0	1	1
0	0	0

EXAMPLE - XOR



XOR		
X1	**X2**	**Y**
1	1	0
1	0	1
0	1	1
0	0	0

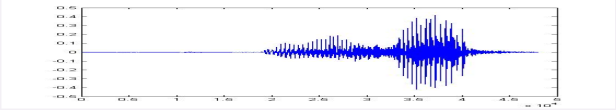
5 MINUTES BREAK

COMBINING NEURONS

DEEP LEARNING OUTLINE

- Introduction of Deep Learning
- “Hello World” for Deep Learning
- Tips for Deep Learning

MACHINE LEARNING ≈ LOOKING FOR A FUNCTION

- Speech Recognition $f($  $)=$ "How are you"
- Image Recognition $f($  $)=$ "Cat"
- Playing Go $f($  $)=$ "5-5" (next move)
- Dialogue System $f($ "Hi" (what the user said) $)=$ "Hello" (system response)

THREE STEPS FOR DEEP LEARNING

Step 1: define a set of function

Neural Network

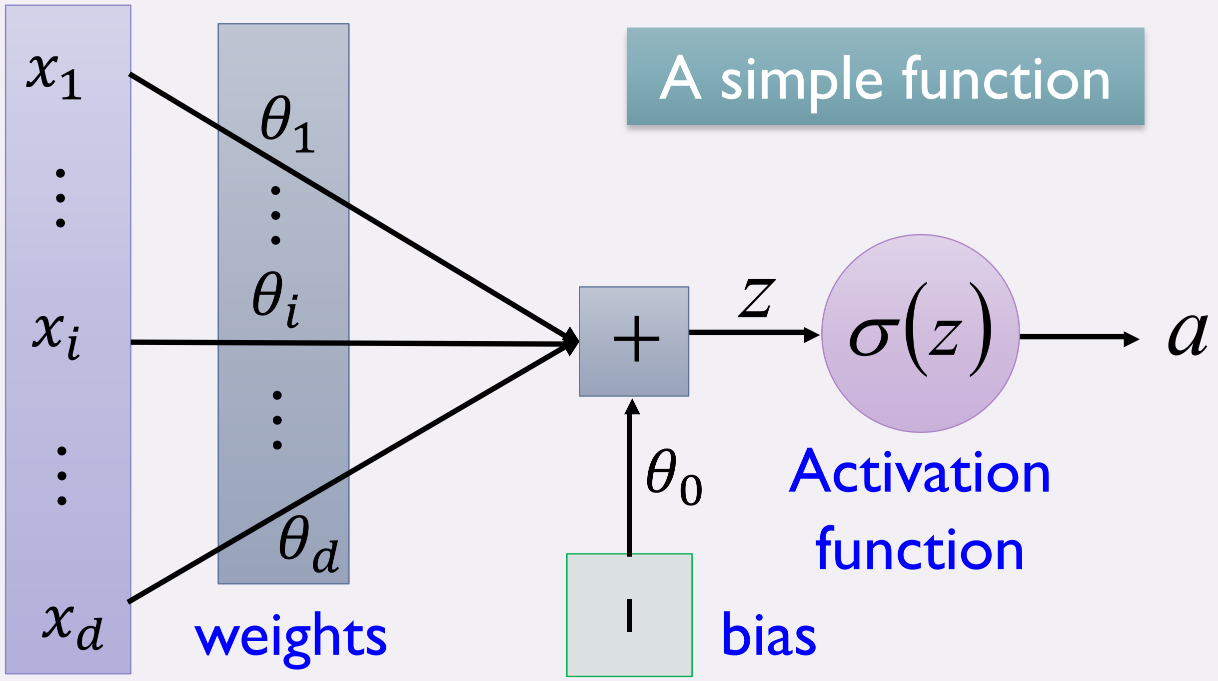
Step 2: goodness of function

Step 3: pick the best function

NEURAL NETWORK

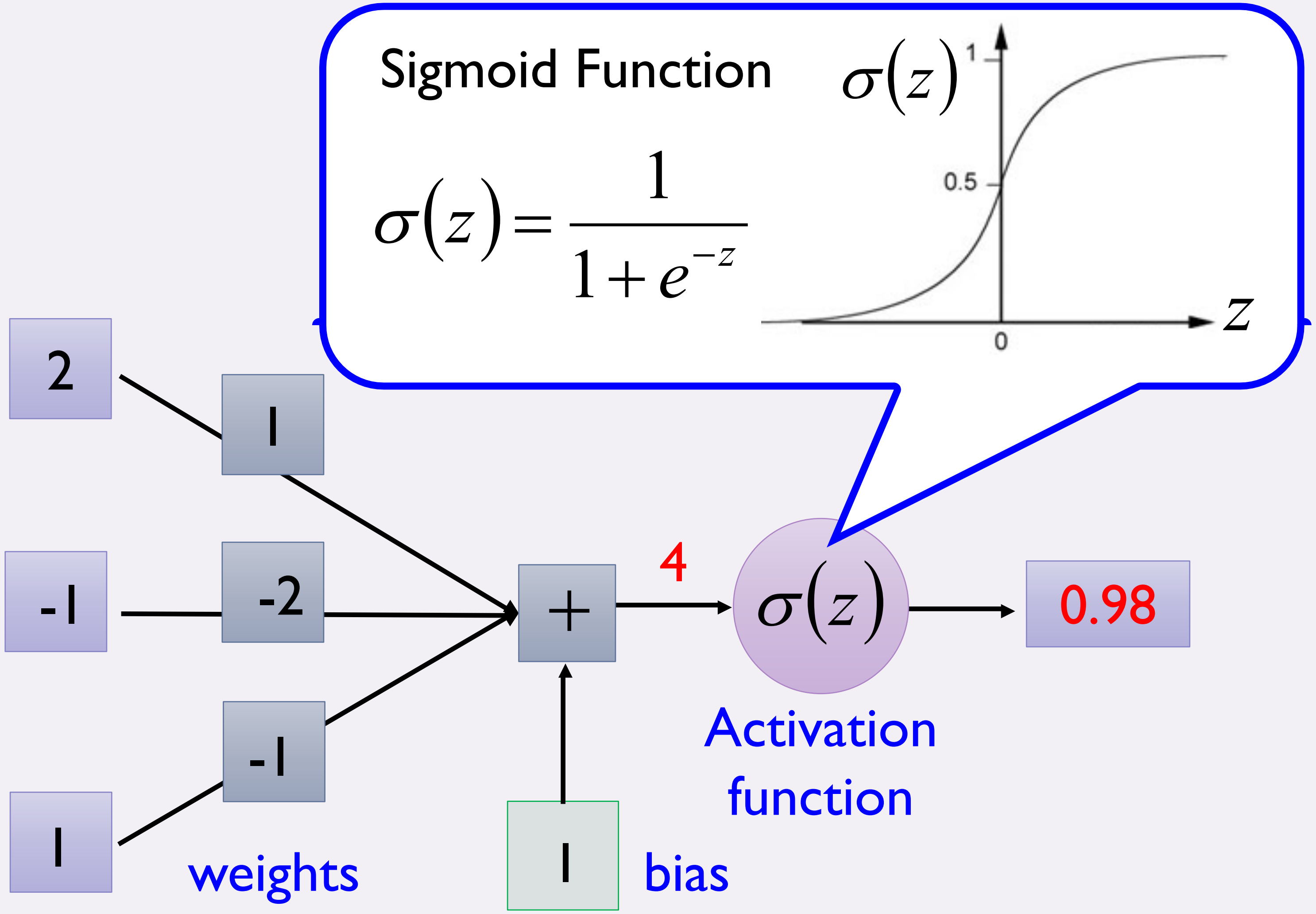
Neuron

$$z = \theta_0 + x_1\theta_1 + \ldots + x_d\theta_d$$

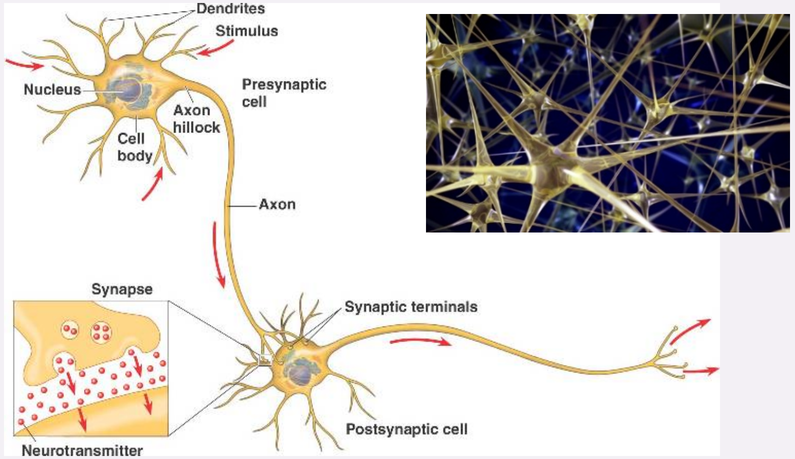


NEURAL NETWORK

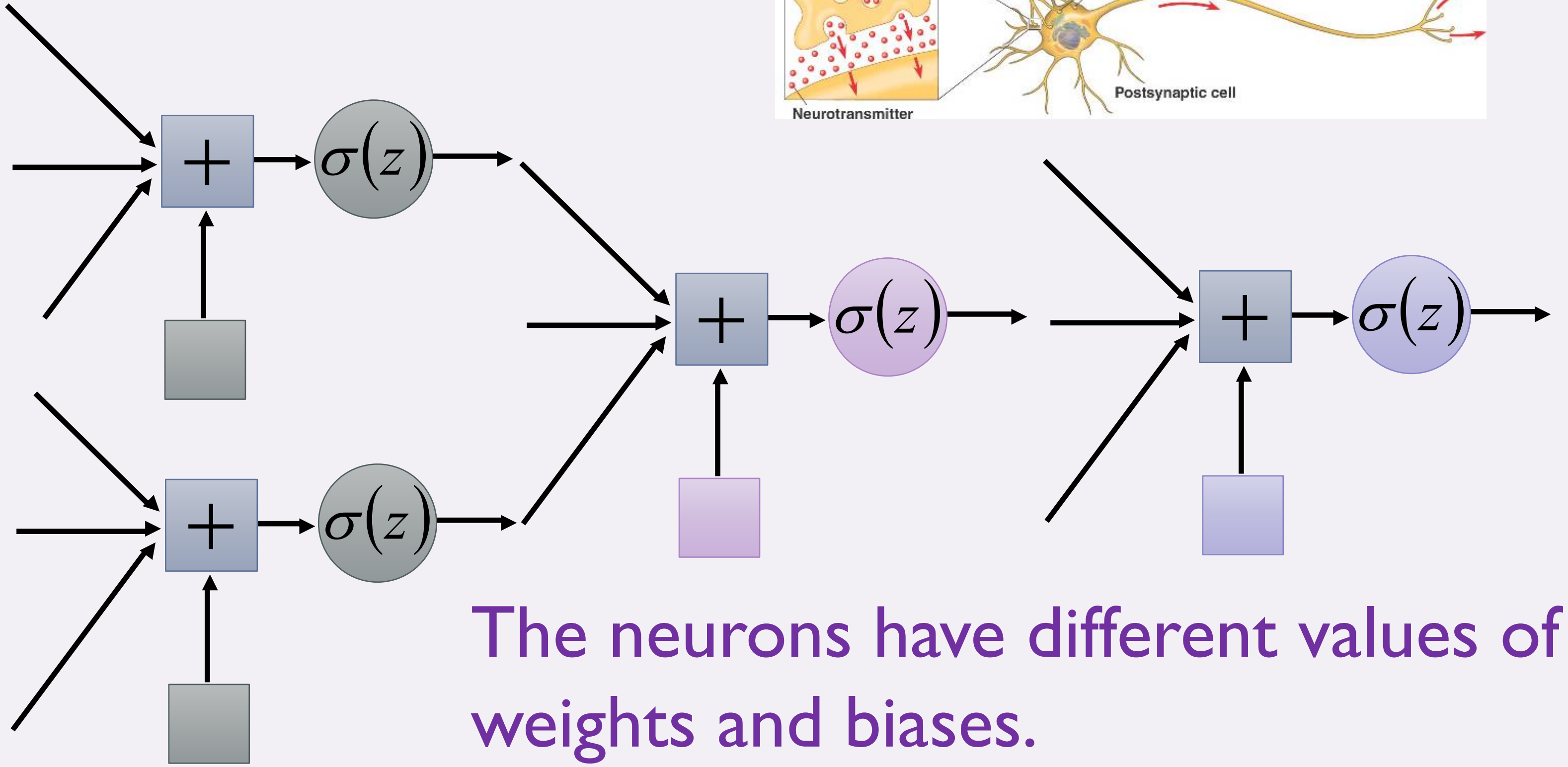
Neuron



NEURAL NETWORK



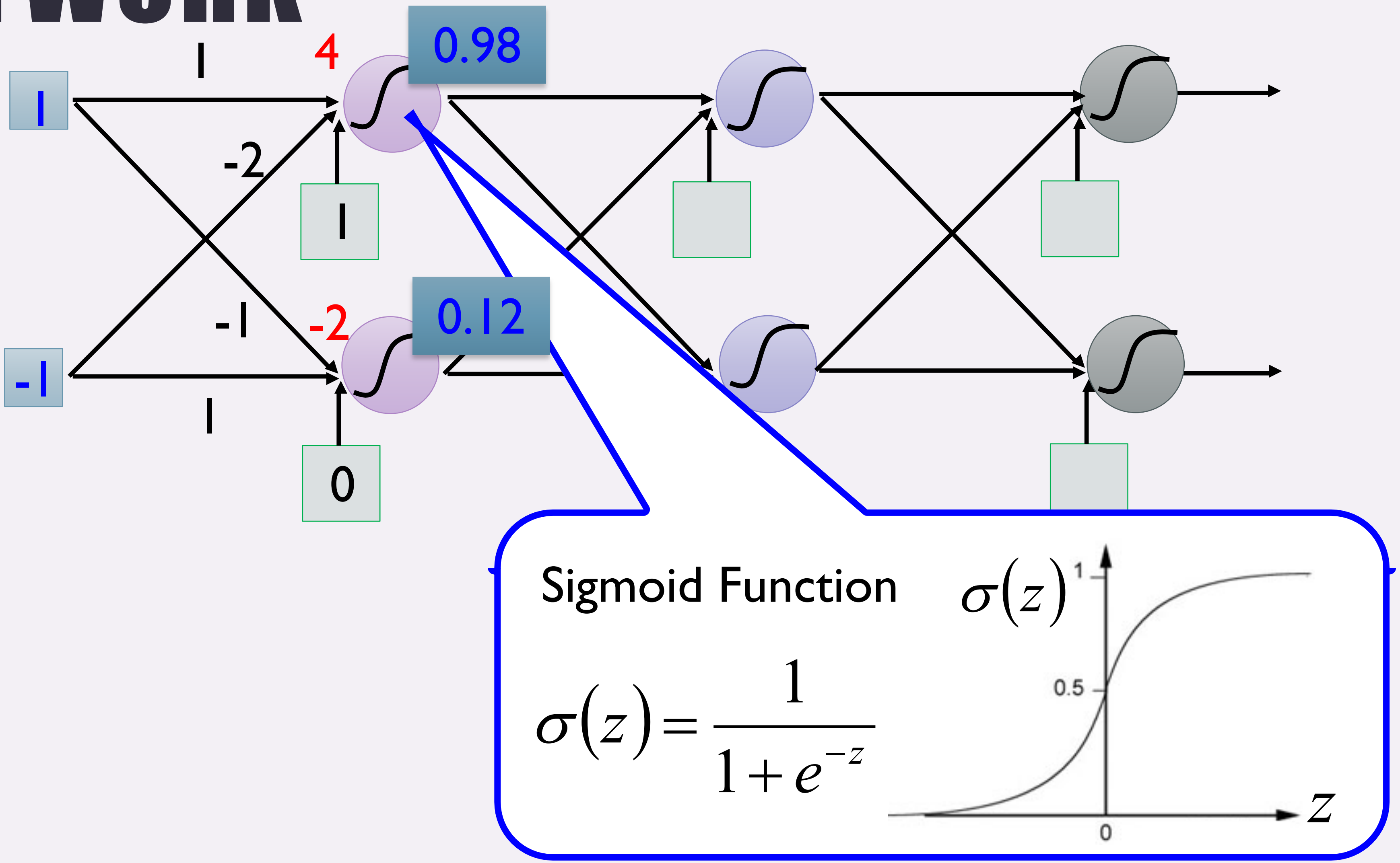
Different connections lead to different network structures



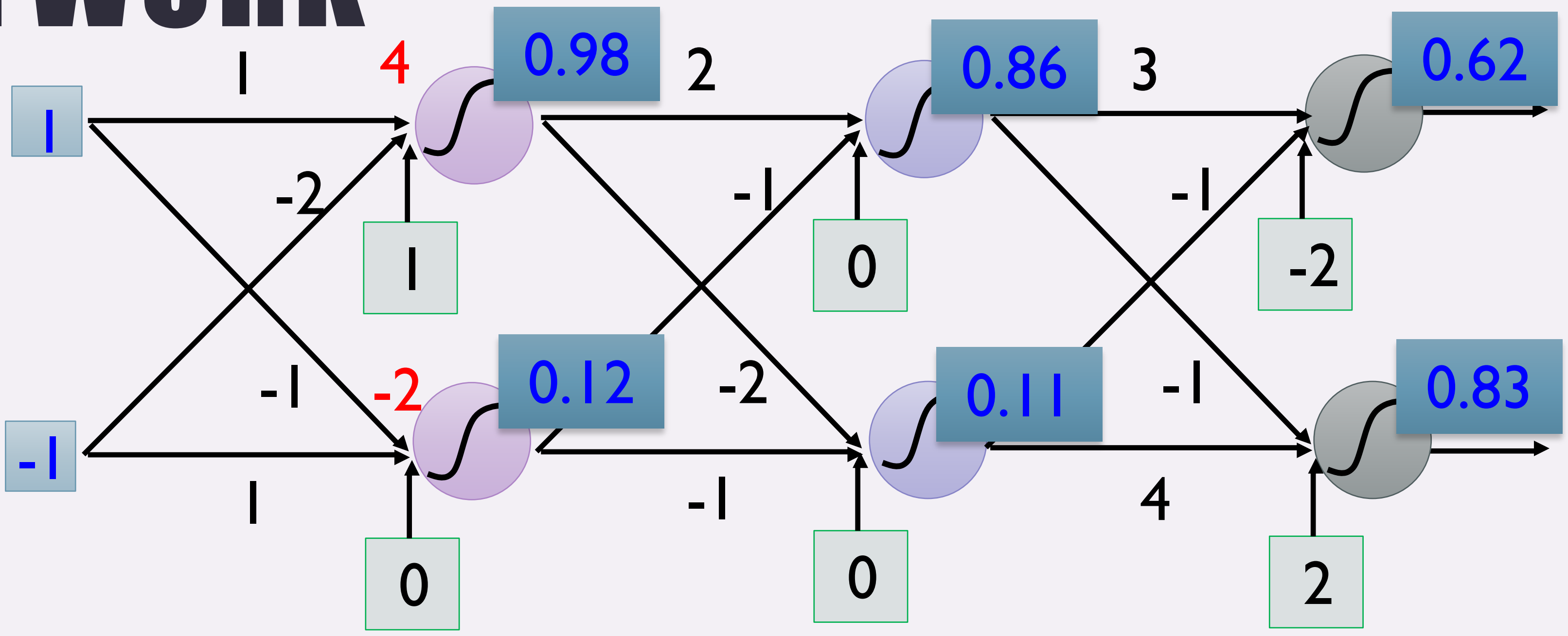
The neurons have different values of weights and biases.

Weights and biases are network parameters θ

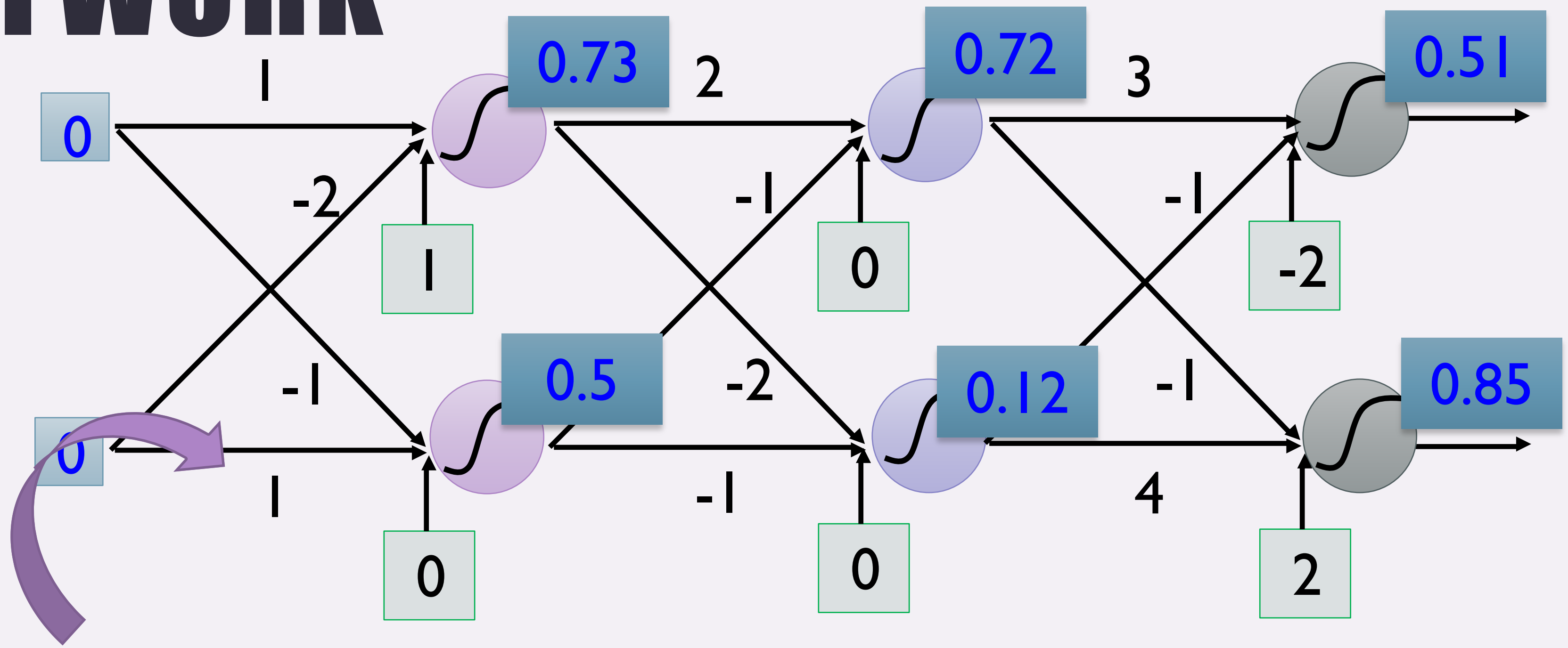
FULLY CONNECT FEEDFORWARD NETWORK



FULLY CONNECT FEEDFORWARD NETWORK



FULLY CONNECT FEEDFORWARD NETWORK



This is a function.

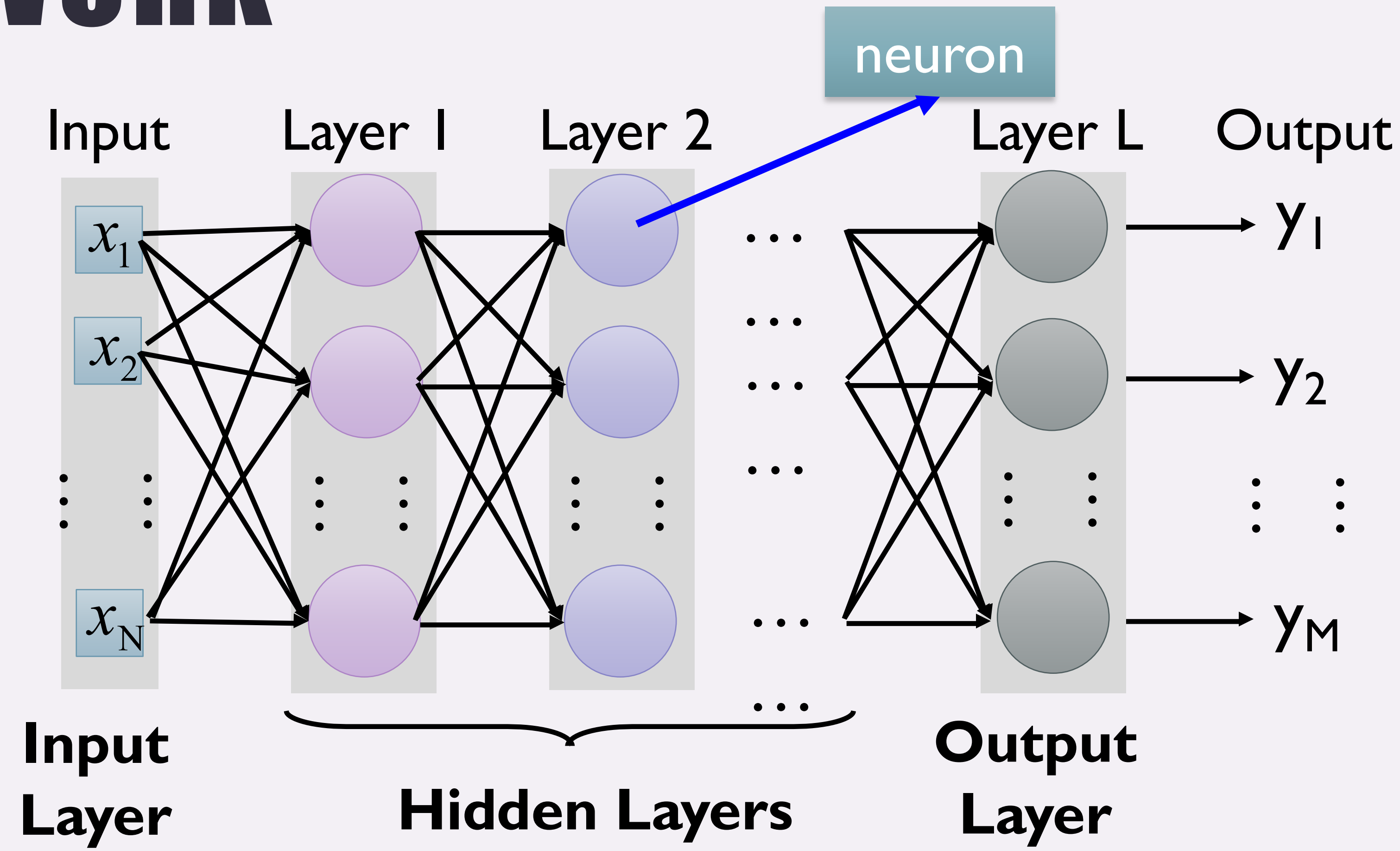
Input vector, output vector

$$f\left(\begin{bmatrix}1\\-1\end{bmatrix}\right)=\begin{bmatrix}0.62\\0.83\end{bmatrix}\quad f\left(\begin{bmatrix}0\\0\end{bmatrix}\right)=\begin{bmatrix}0.51\\0.85\end{bmatrix}$$

Given parameters θ, define a function

Given network structure, define ***a function set***

FULLY CONNECT FEEDFORWARD NETWORK

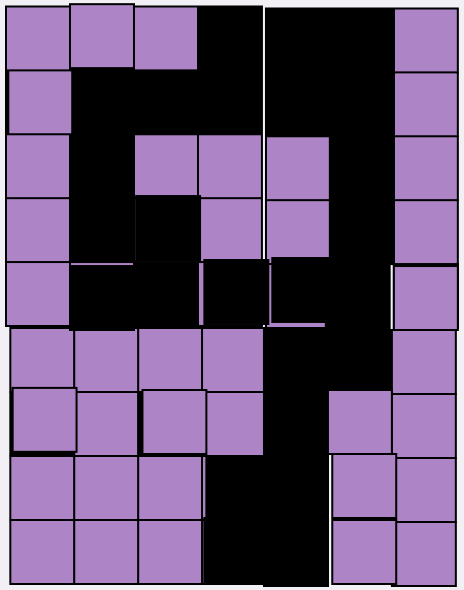


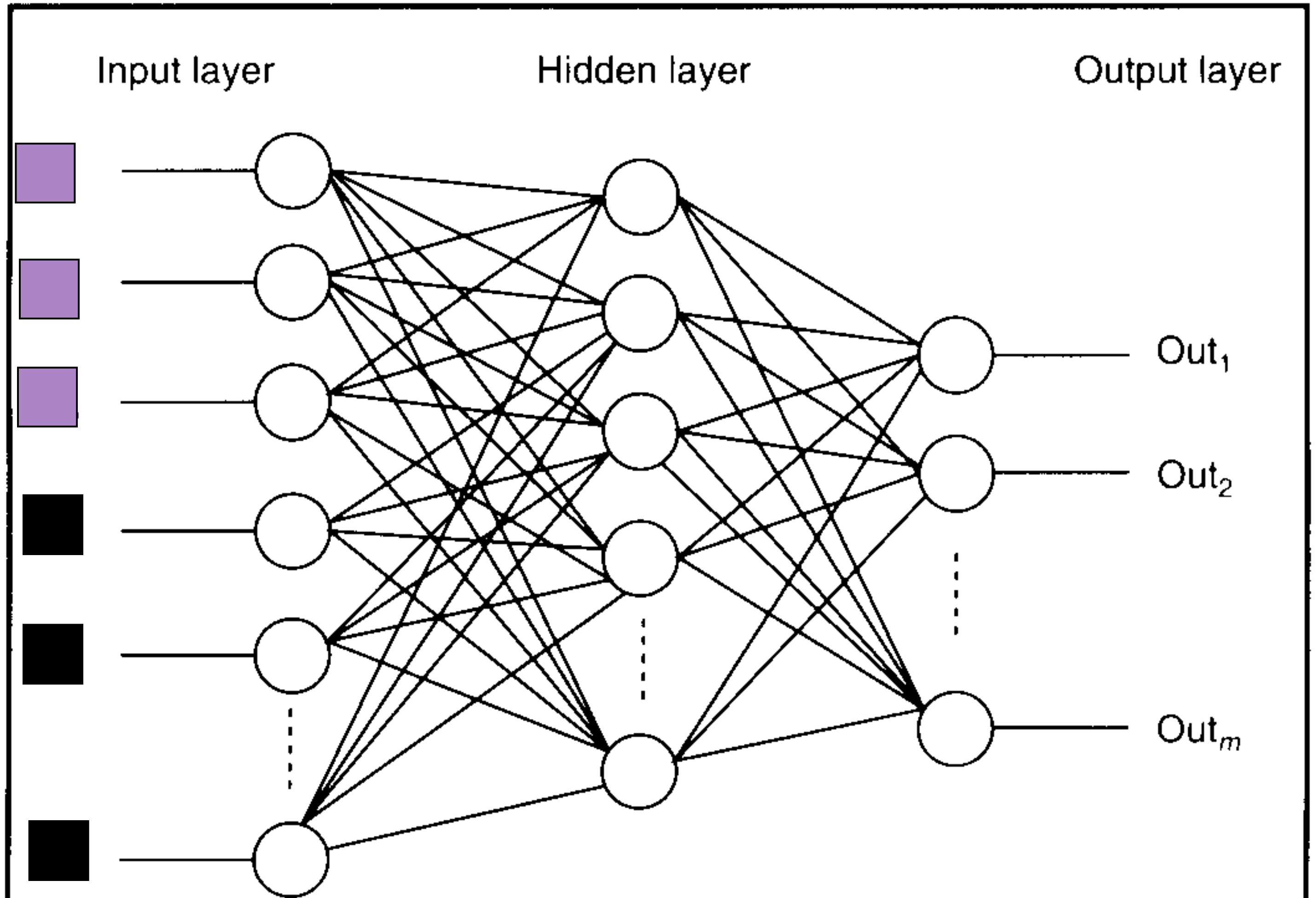
Deep means many hidden layers

EXAMPLE: DIGIT CLASSIFICATION

FEATURE DETECTORS

Figure 1.2: *Examples of handwritten digits from postal envelopes.*





WHAT IS THIS UNIT DOING?

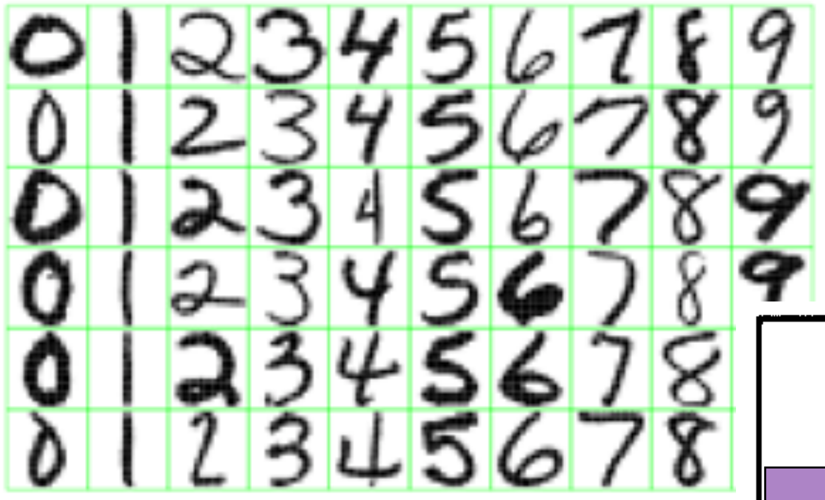
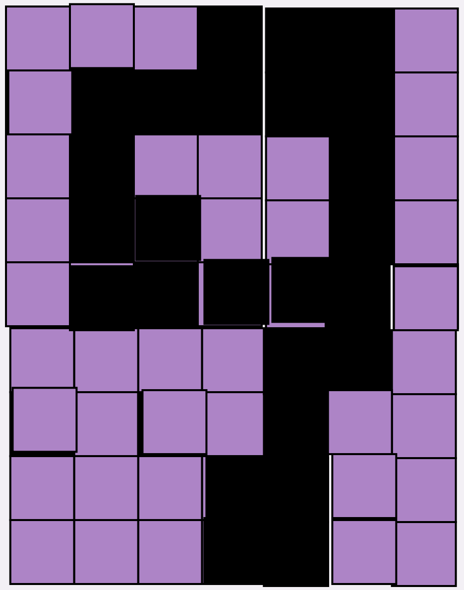
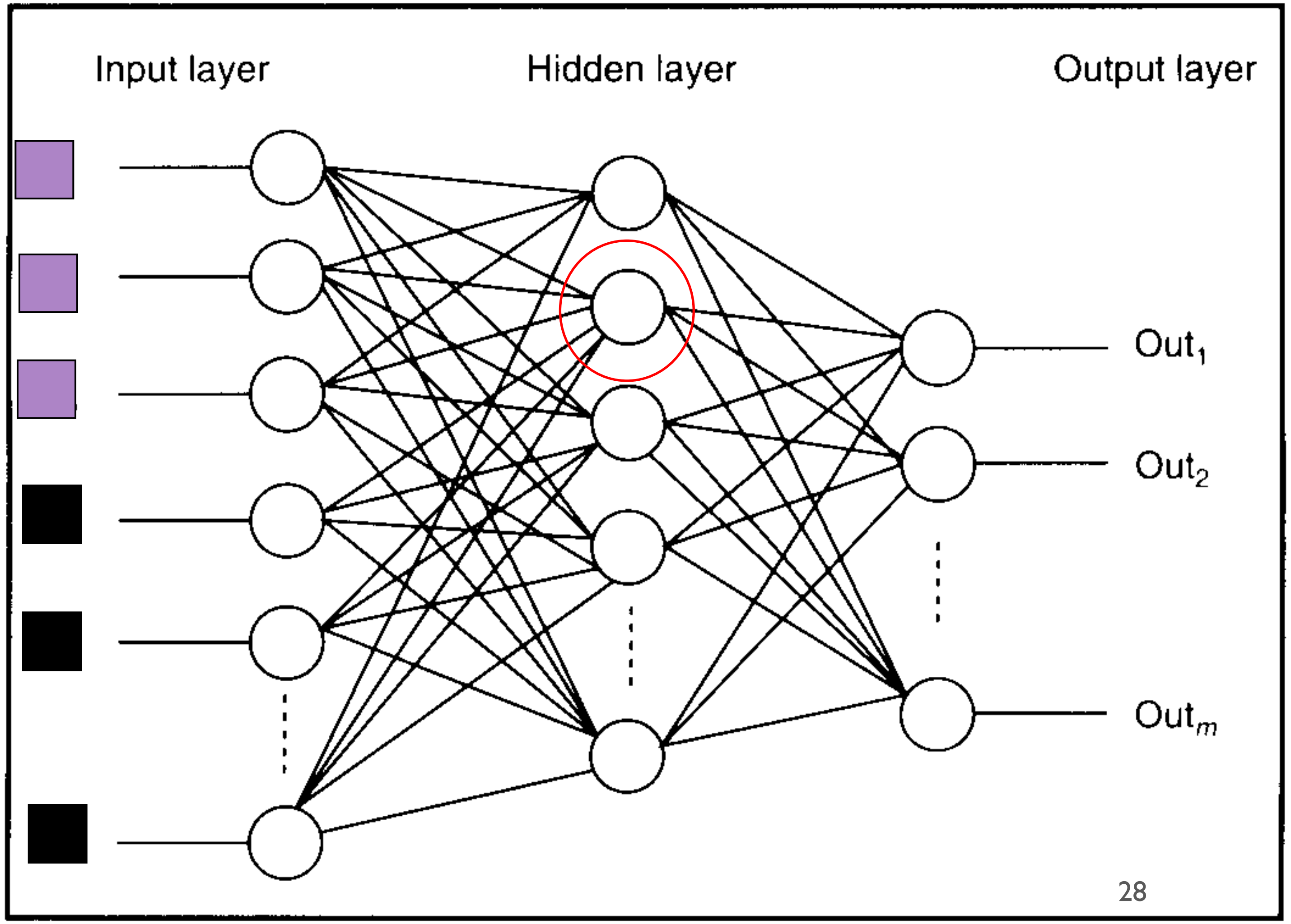
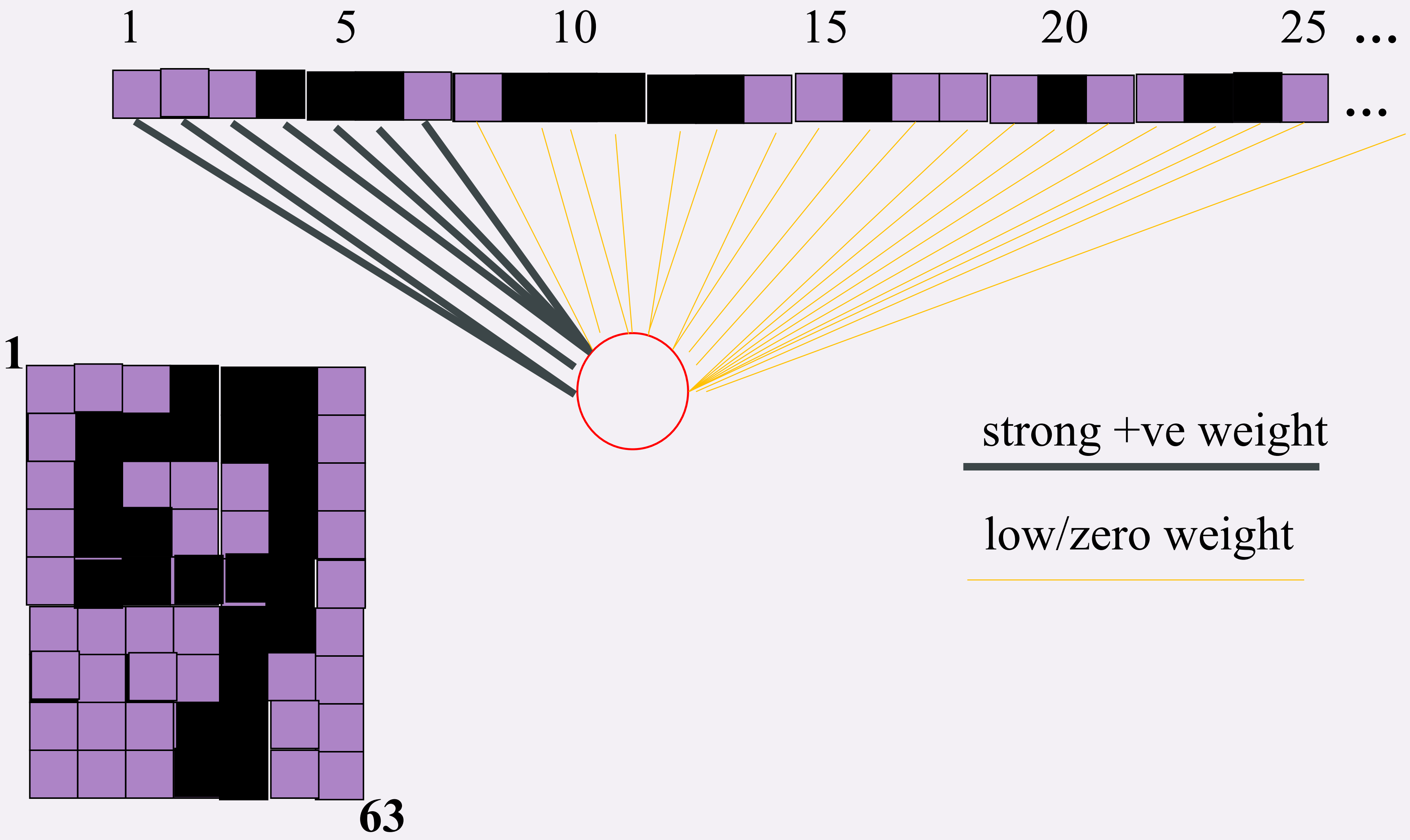


Figure 1.2: *Examples of handwritten digits from postal envelopes.*

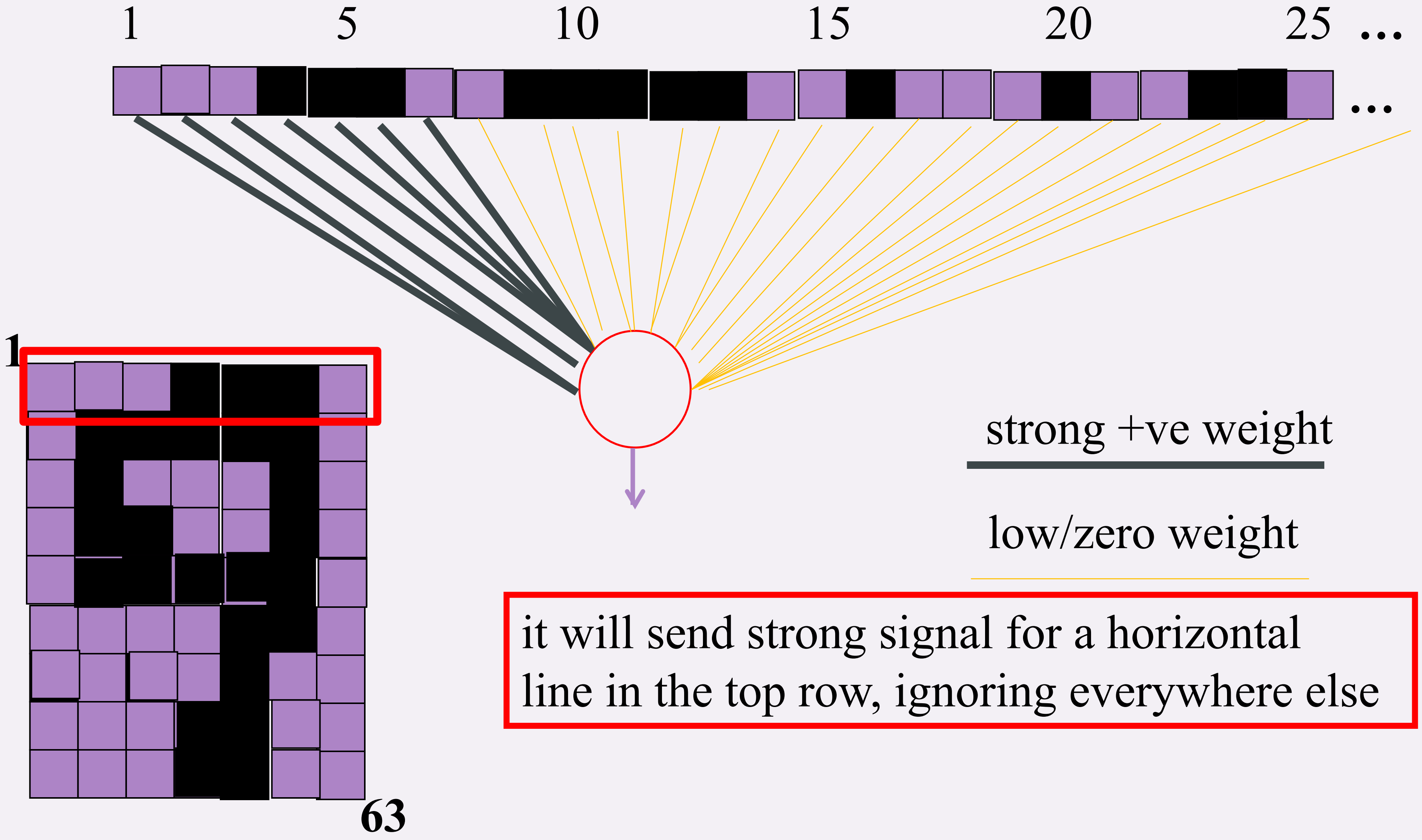




SELF-ORGANIZED FEATURE DETECTORS

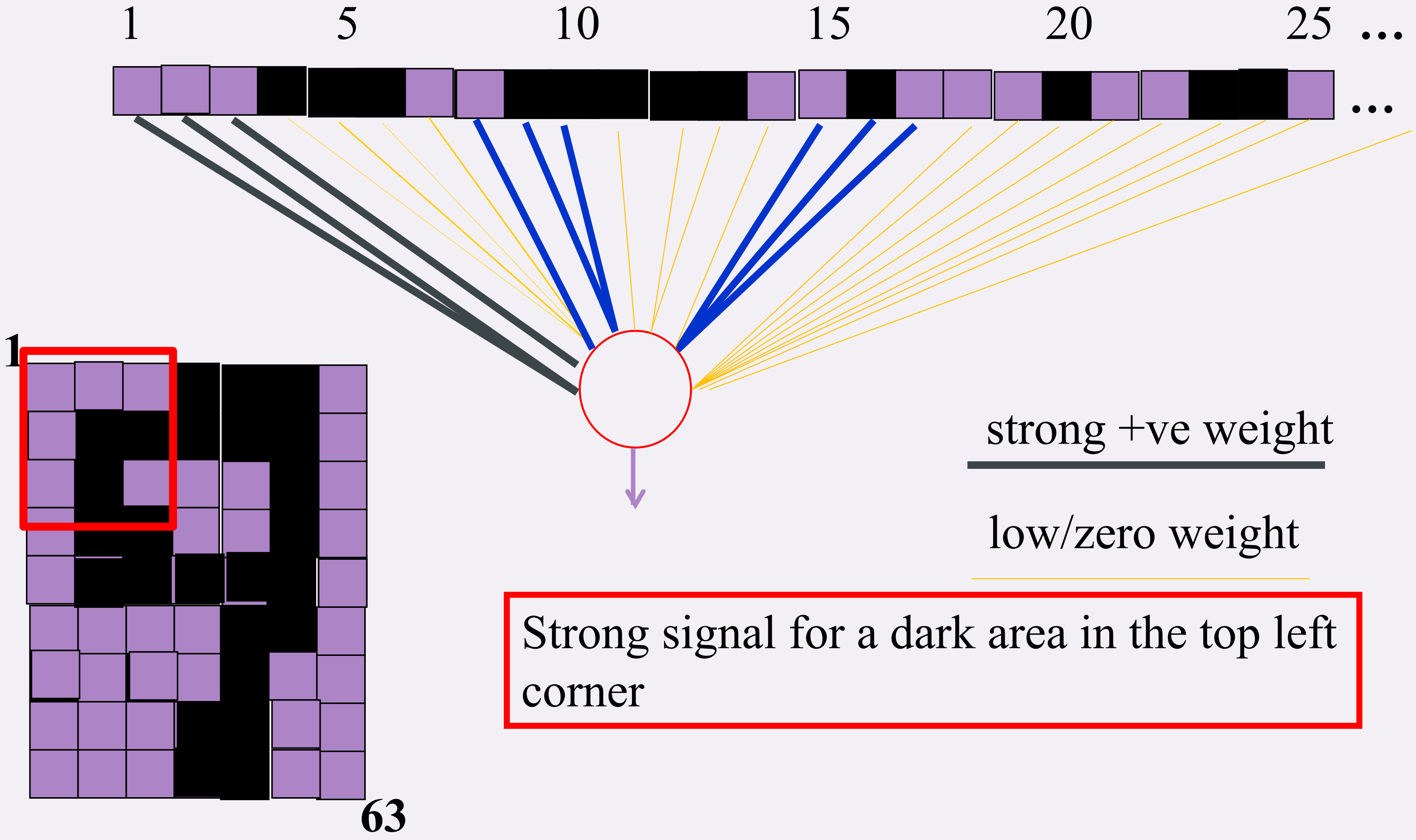


WHAT DOES THIS UNIT DETECT?

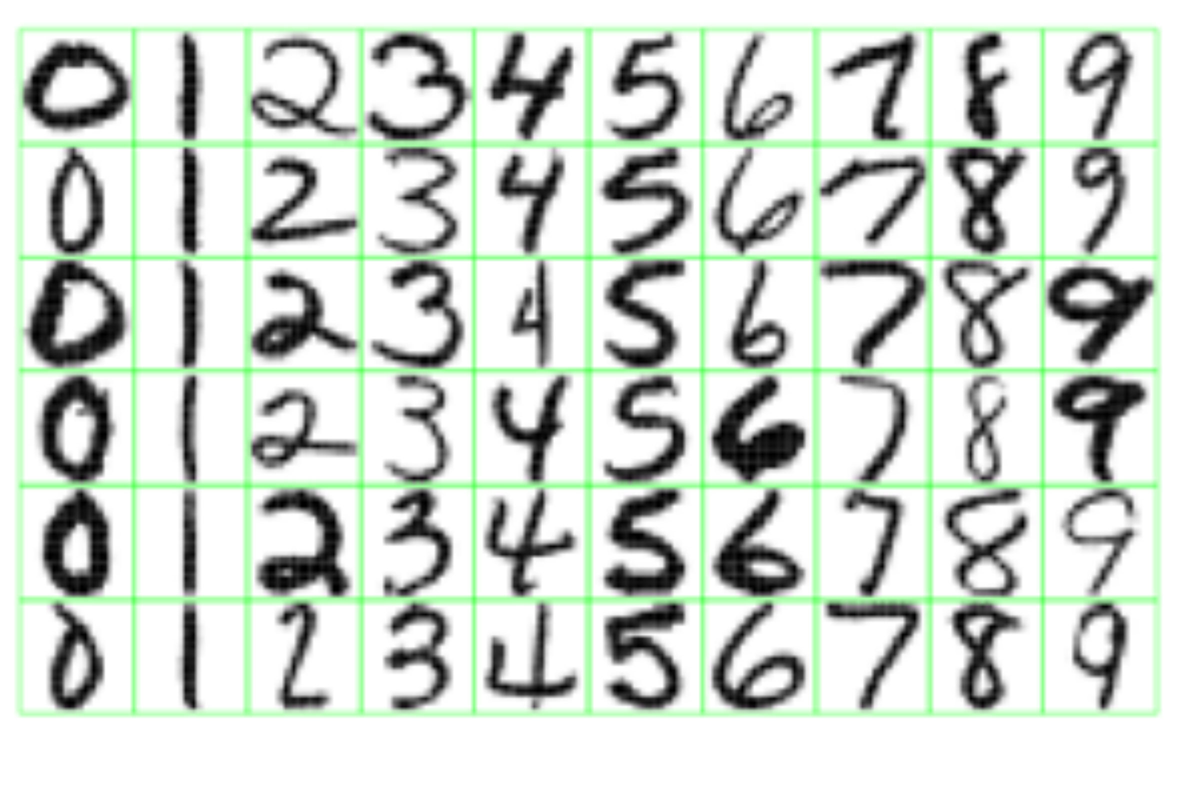


it will send strong signal for a horizontal line in the top row, ignoring everywhere else

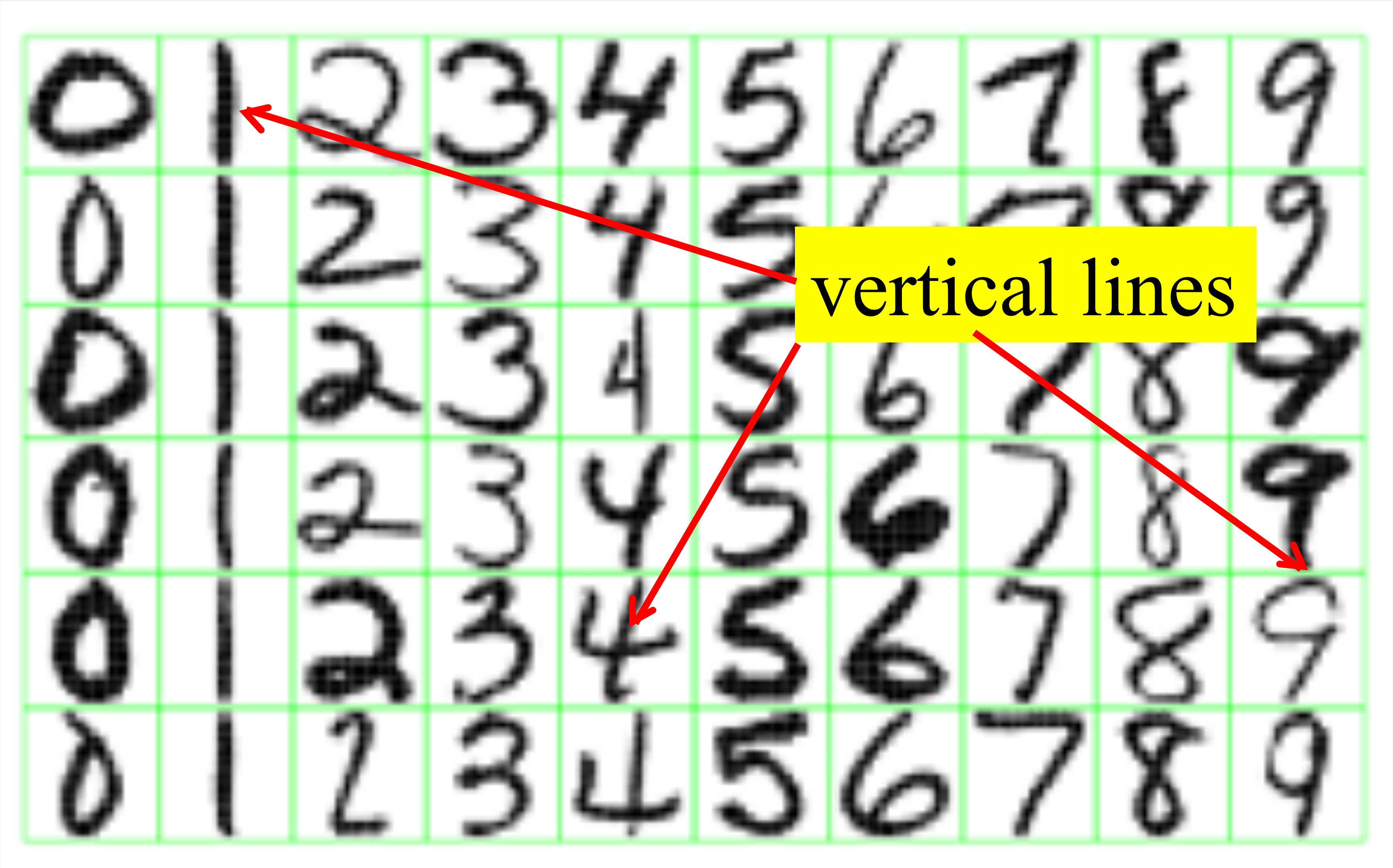
WHAT DOES THIS UNIT DETECT?



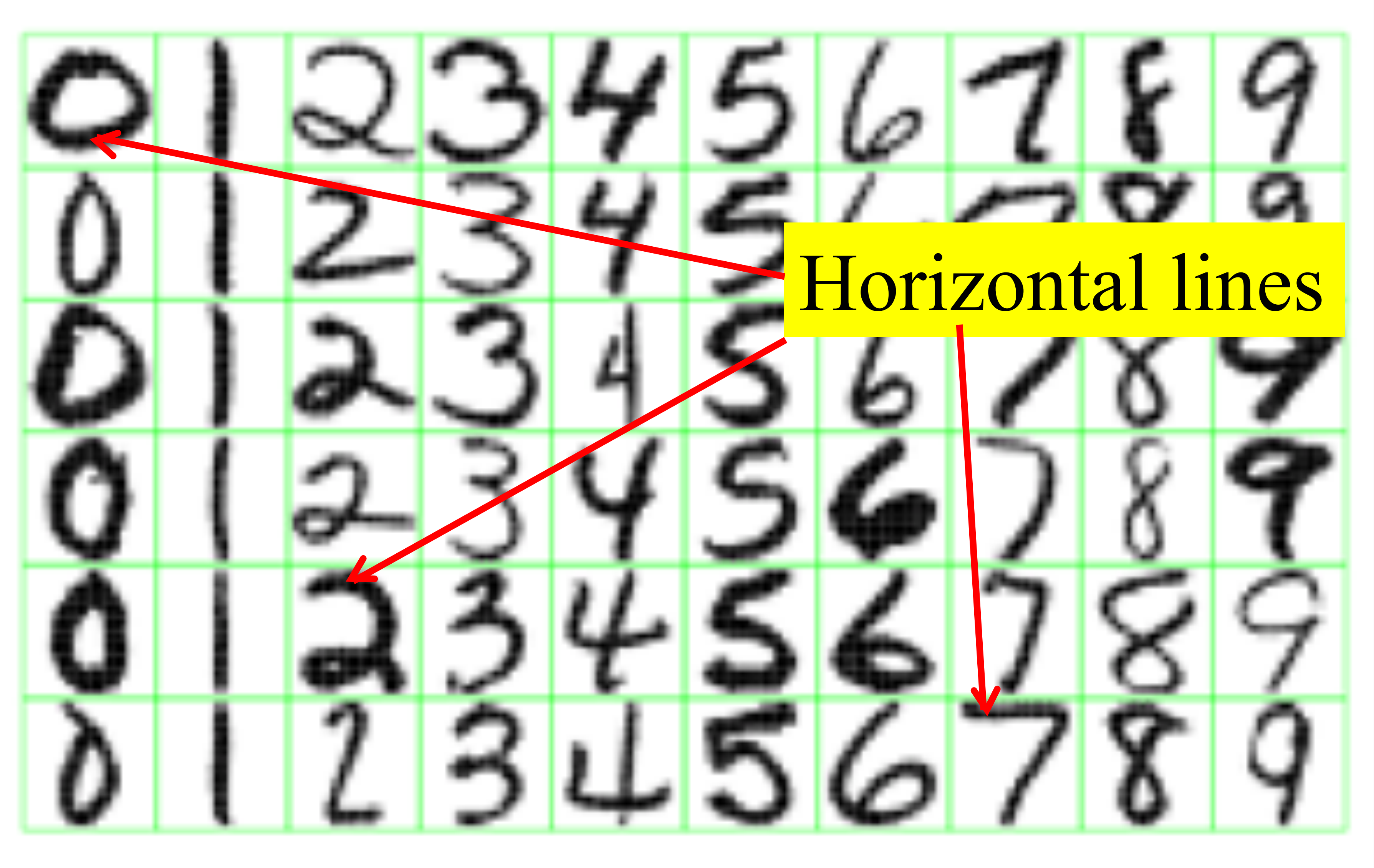
Strong signal for a dark area in the top left corner

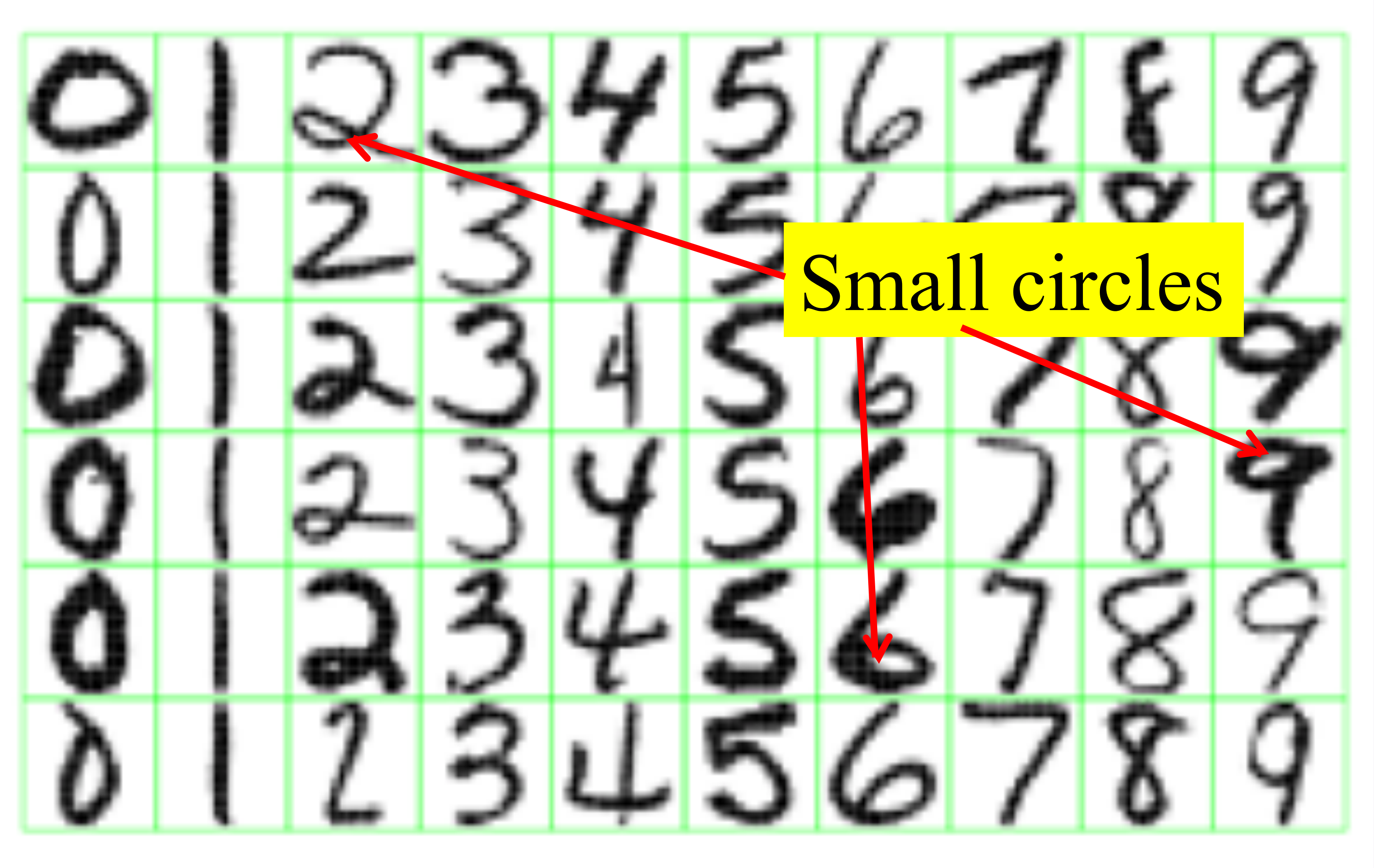


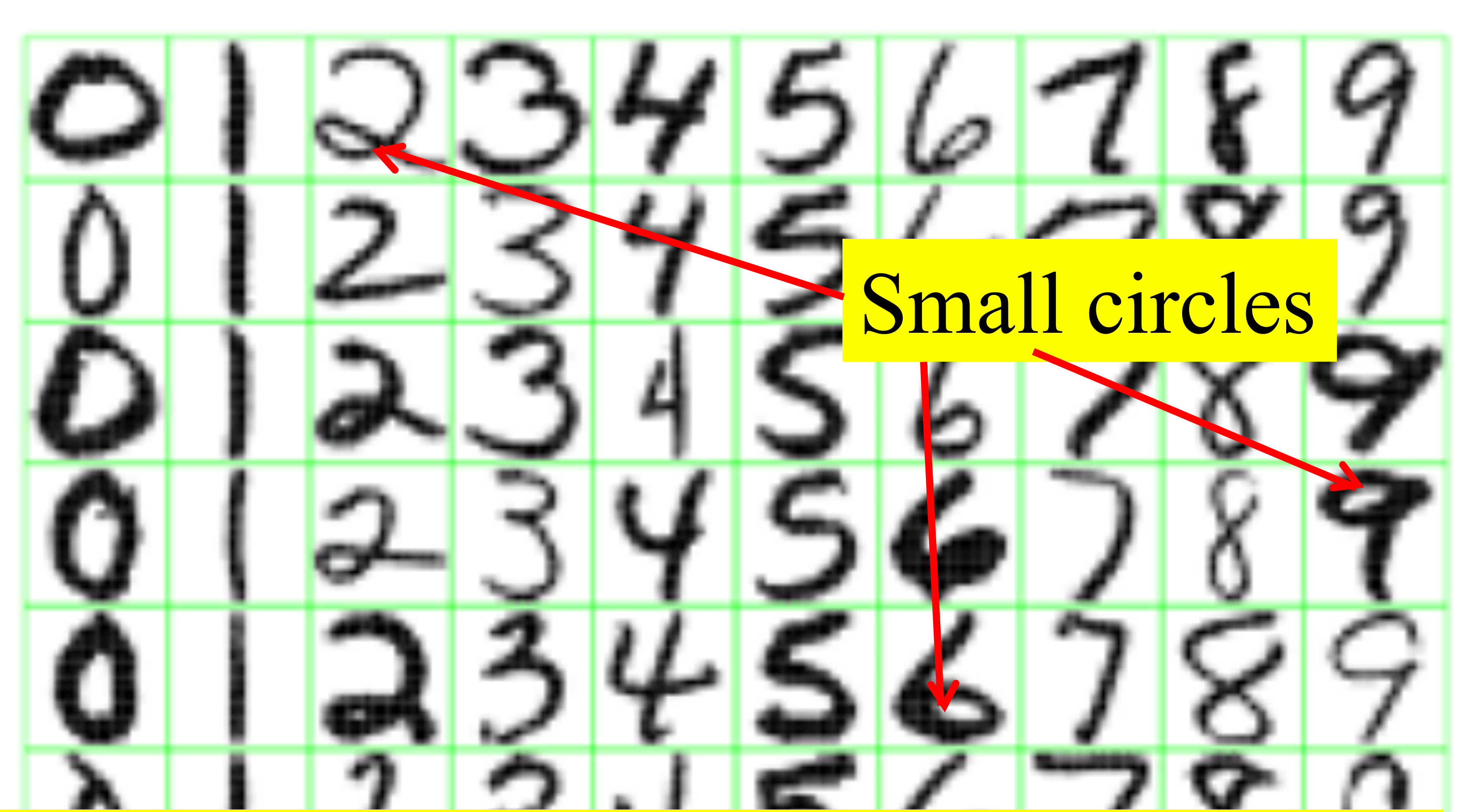
What features might you expect a good NN to learn, when trained with data like this?



1

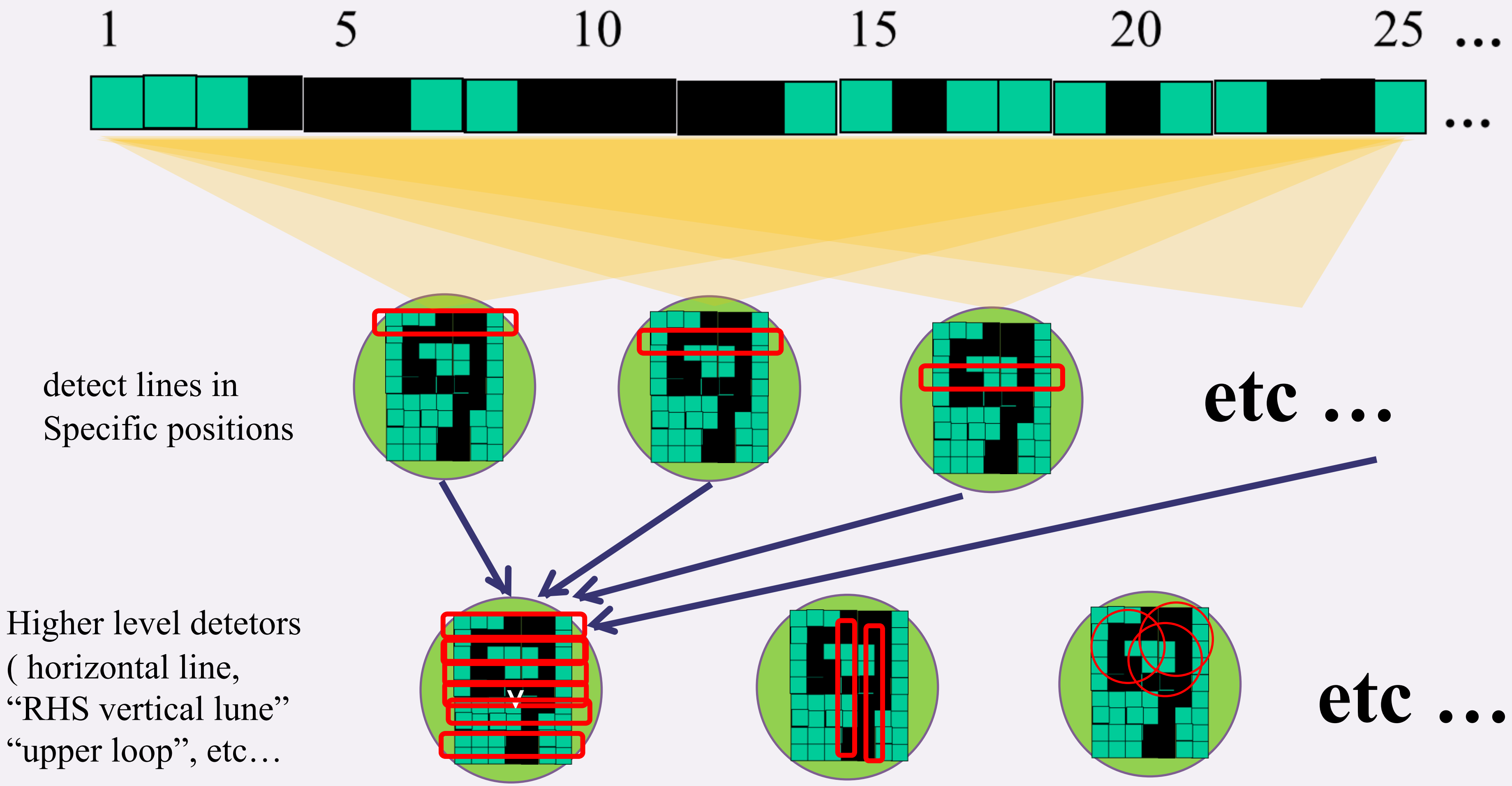

Horizontal lines


Small circles

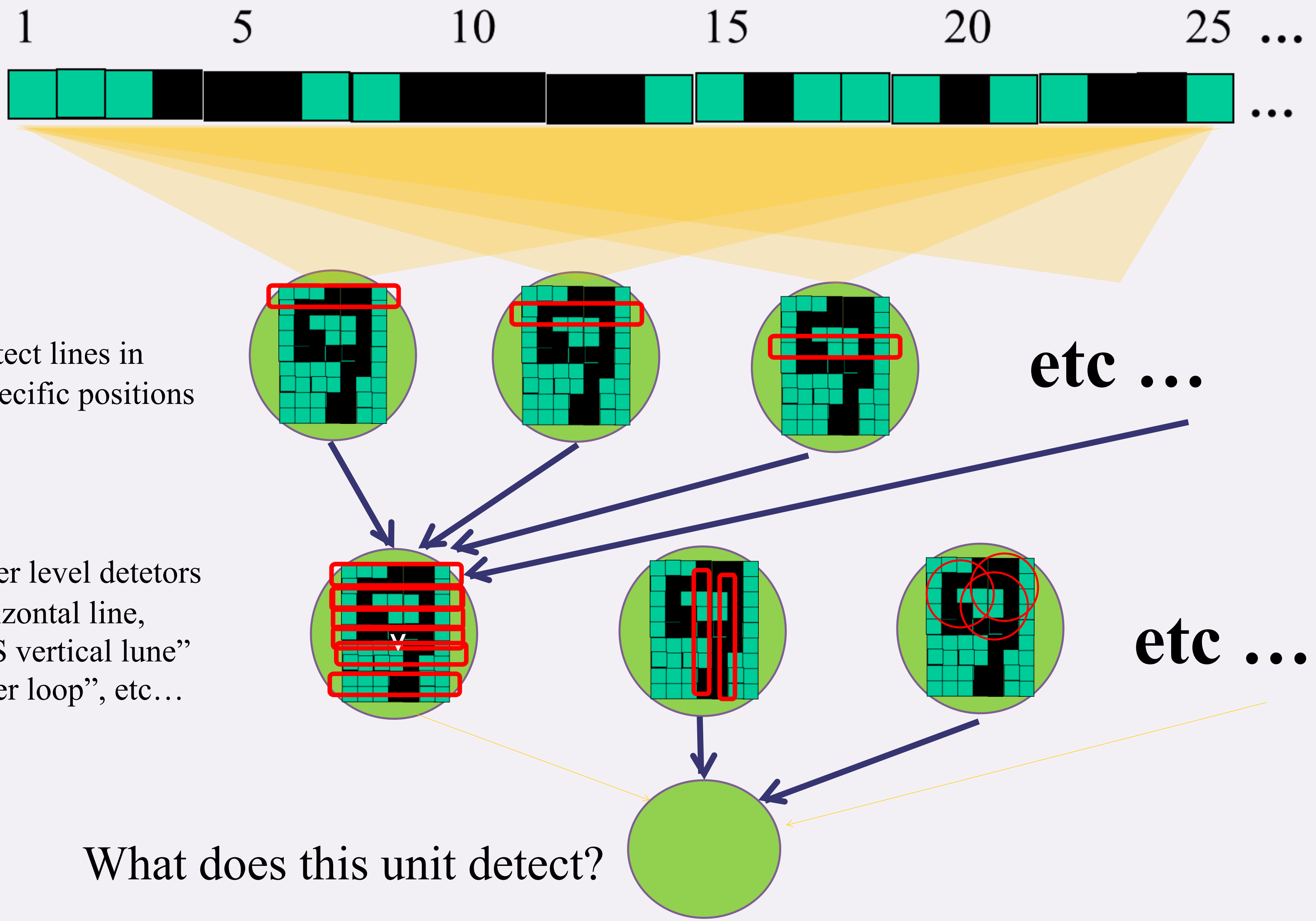


But what about position invariance ???
our example unit detectors were tied to specific parts of the image

SUCCESSIVE LAYERS CAN LEARN HIGHER-LEVEL FEATURES …



SUCCESSIVE LAYERS CAN LEARN HIGHER-LEVEL FEATURES …



detect lines in
Specific positions

Higher level detetors
(horizontal line,
“RHS vertical lune”
“upper loop”, etc…

What does this unit detect?

EXERCISE

HTTPS://BIT.LY/3EK0SPH

QUESTIONS?