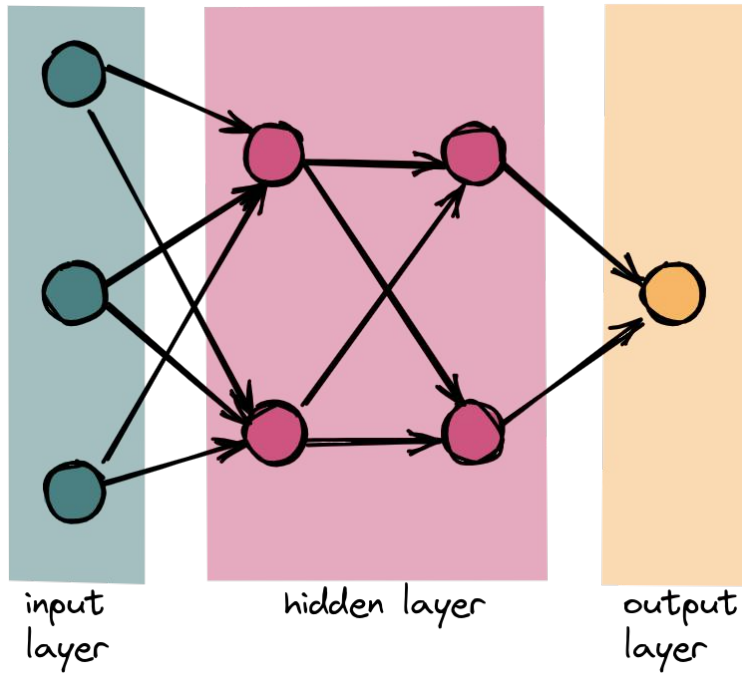# Neural Networks
Part 1

Week 15

Middlesex University Dubai; CST4050 Fall21;
Instructor: Ivan Reznikov

Plan

- Neural networks terminology

- Perceptron

- Activation functions

- Neural network: example case

What methods we've covered?

1. Linear Regression
2. Logistic regression
3. Decision Trees and Random Forest
4. Support Vector Machines
5. Naive Bayes

Neural Networks

Neural networks are an extremely popular ML technique.

Their popularity is the result of high-quality predictions neural networks make.

Methods studied previously allowed us to find N-dimensional dependencies where N wasn't significant. But for most neural networks finding complex fitting curves is possible.
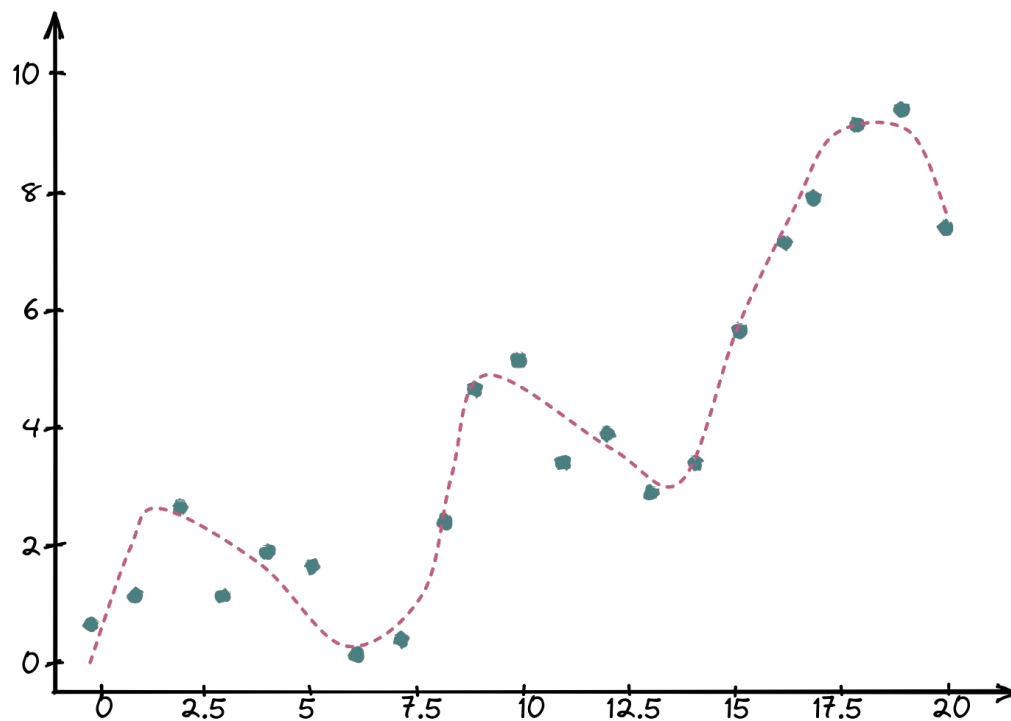
Neural Networks

Modern neural networks have nothing to do with biology. It's more like the "broken phone" game. But in neural networks, transforming information leads to better results, though.

Neural networks consist of an input layer, one or more hidden layers, and one output layer.

A neural network consists of nodes and connections.

nodes

input layer

hidden layer

output layer

connections

5
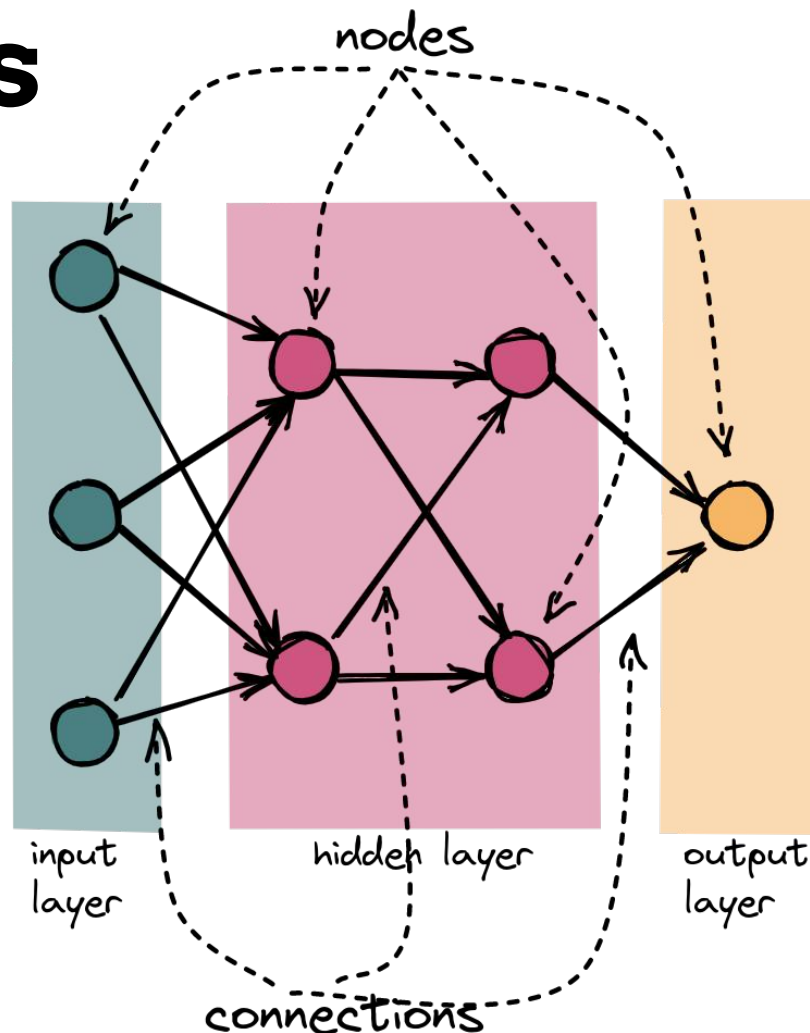
Perceptron

A **perceptron** is a simple unit that takes several inputs, aggregates them, and passes the result through a function to give a single result.

The aggregation is usually the combination of linear equations for different node values. Each child node has its "**weight**," which is summed up with the parents' "**bias**."

As the equation describes the line, we can say perceptron is a linear classifier.
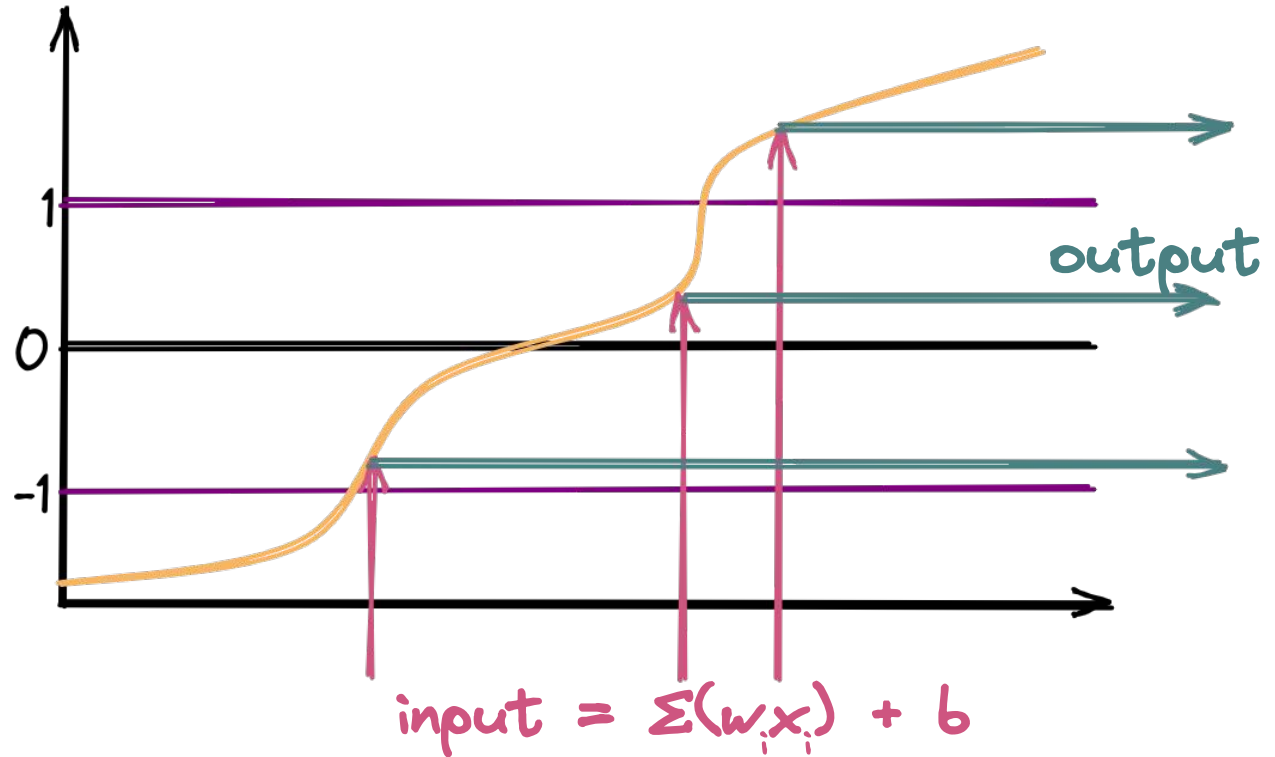
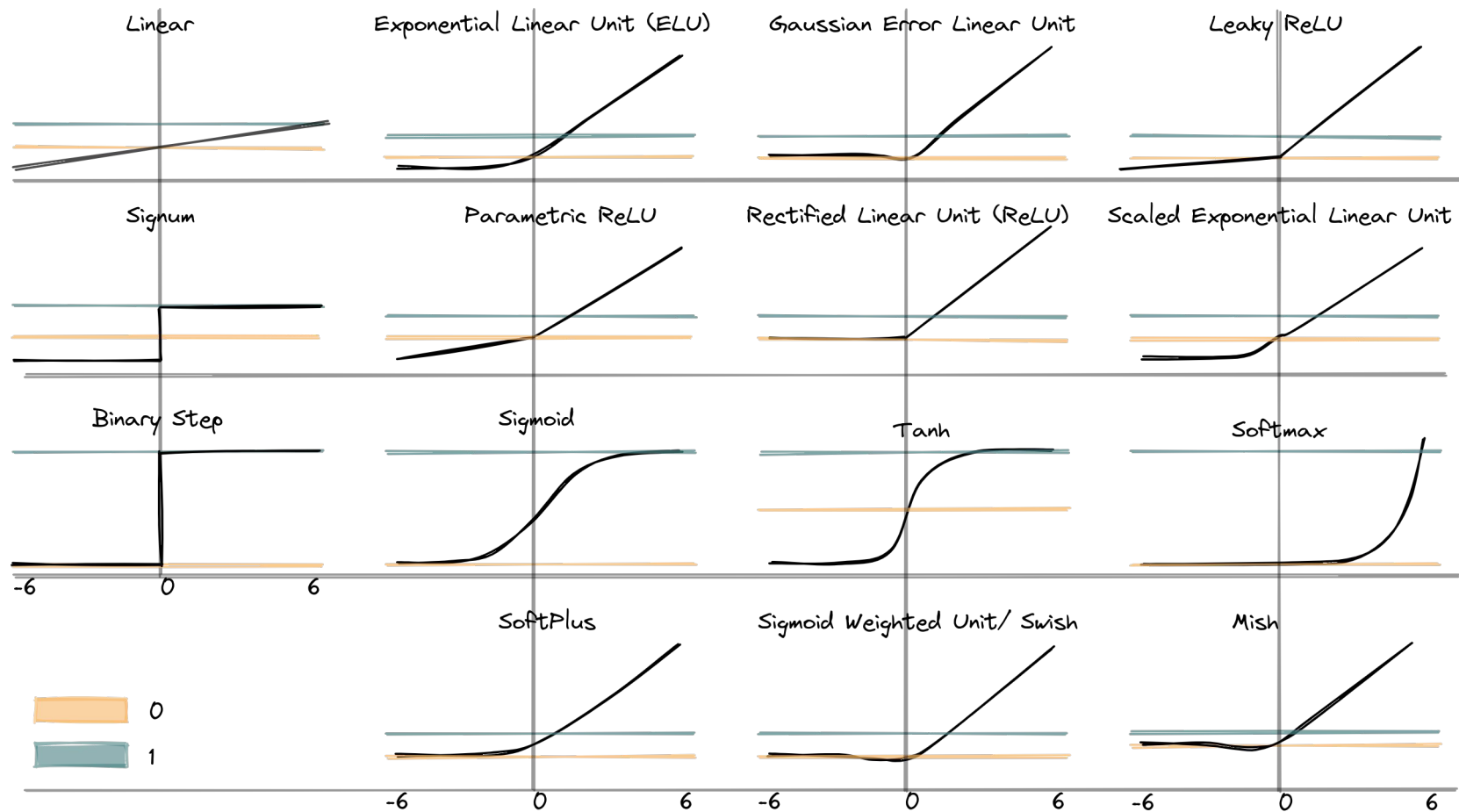$$y = wx + b$$
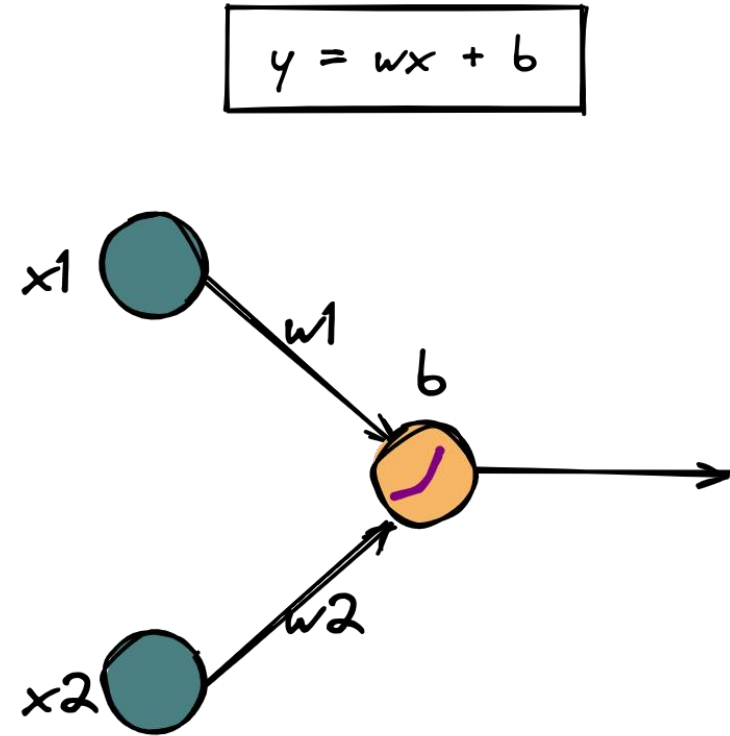
$$y = w1x1 + w2x2 + b$$

Activation function

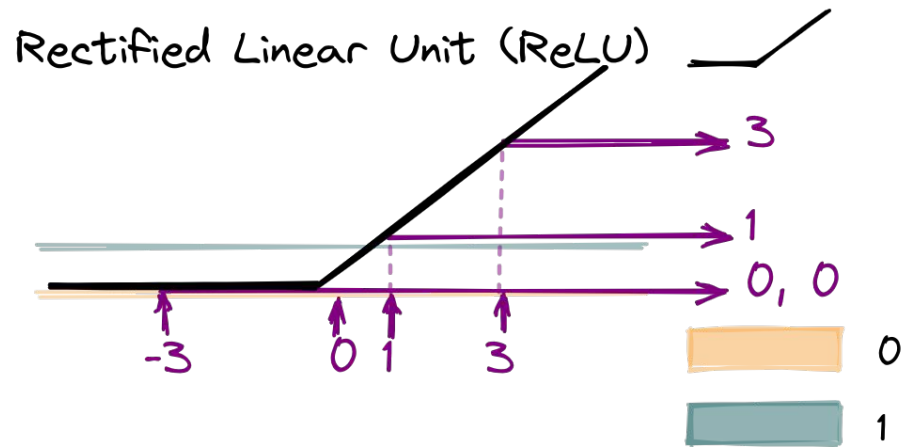An activation function is the key for a neural network to learn complex patterns in the data.

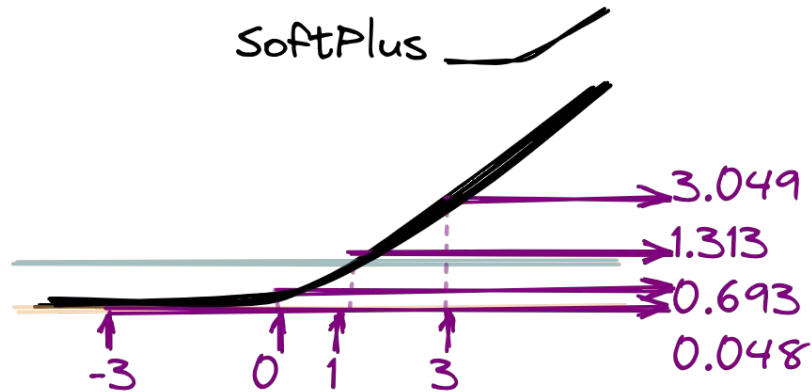The most crucial ability of activation functions is adding non-linearity to a neural network.

output

$$input = \Sigma(w_i x_i) + b$$

Activation functions

Activation functions: examples

Binary Step

1, 1
0, 0
-3   0 1   3

Sigmoid

0.953
0.731
0.5
0.047
-3   0 1   3

SoftPlus

3.049
1.313
0.693
0.048
-3   0 1   3

Rectified Linear Unit (ReLU)

3
1
0, 0
-3   0 1   3

0
1
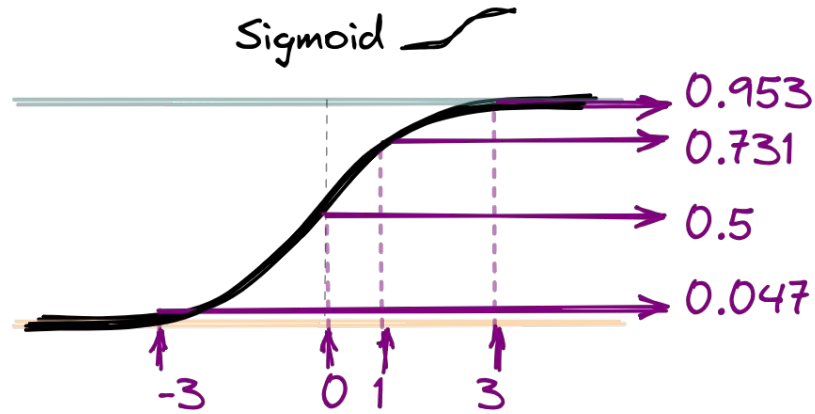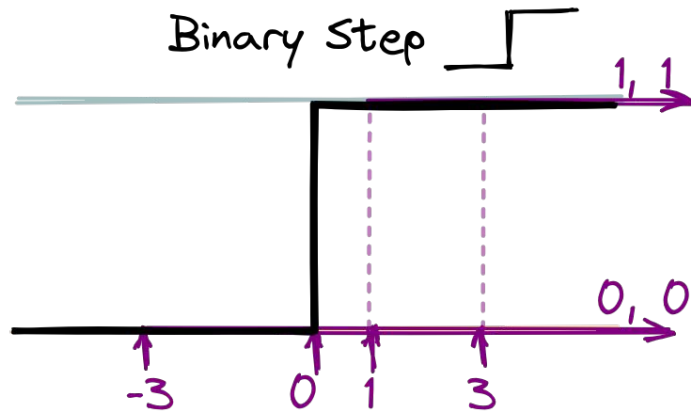
Neural Networks: Case1

For our example we'll use a neural network with one input node, two nodes in the only hidden layer and one output node.

To experiment with different activation functions, we'll use ReLU (-1), sigmoid and softplus.

Weights and biases are usually set randomly in the first epoch.
In our case, we'll use somehow optimized values.

$i \times w1 + b1$

$i_{1,1} \times w3 + b3$

$i \times w2 + b2$

$i_{1,2} \times w4 + b4$

input   hidden   output
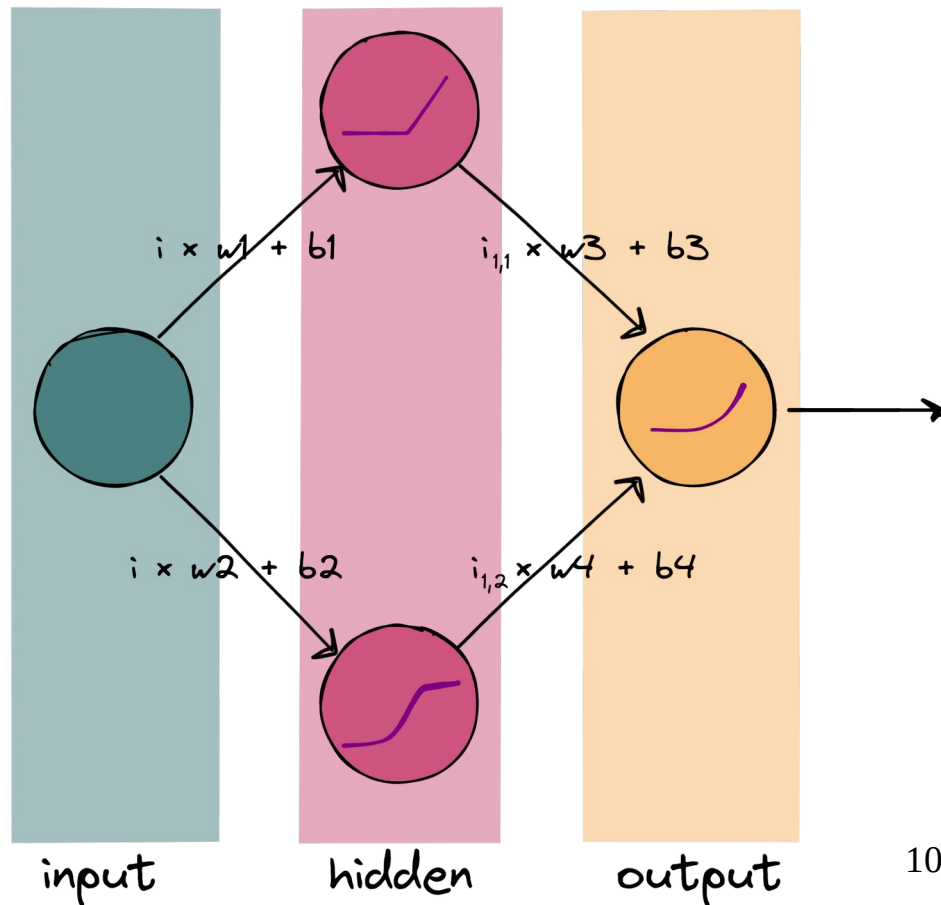
Case 1: Data Visualization

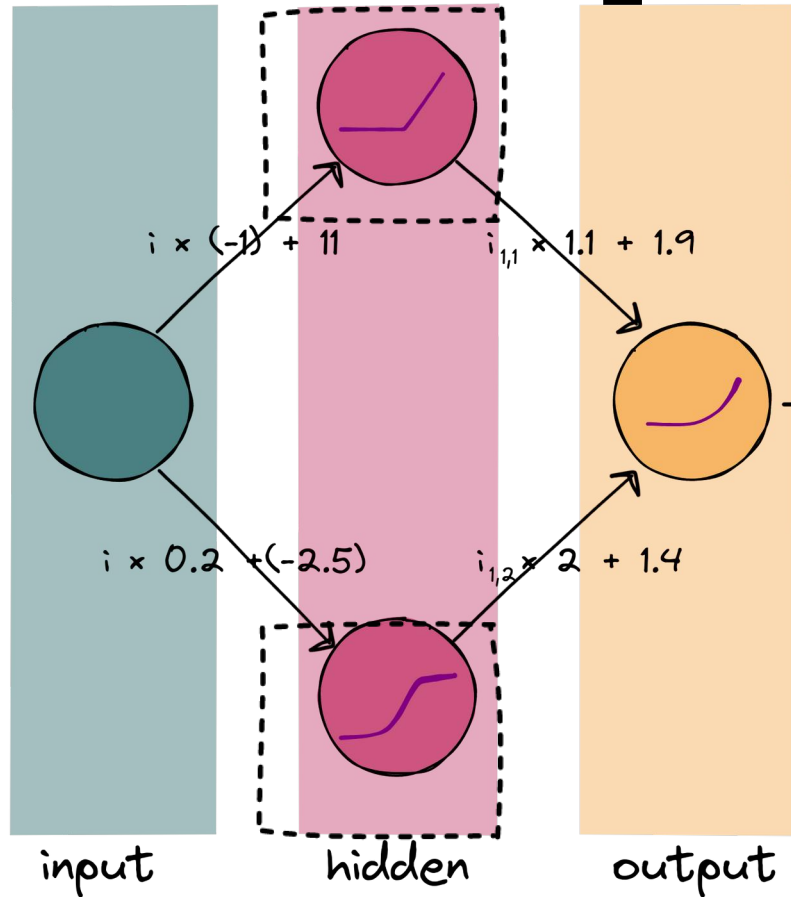| x	y
10	5.19
11	3.38
12	3.78
13	2.78
14	3.31
15	5.69
16	7.15
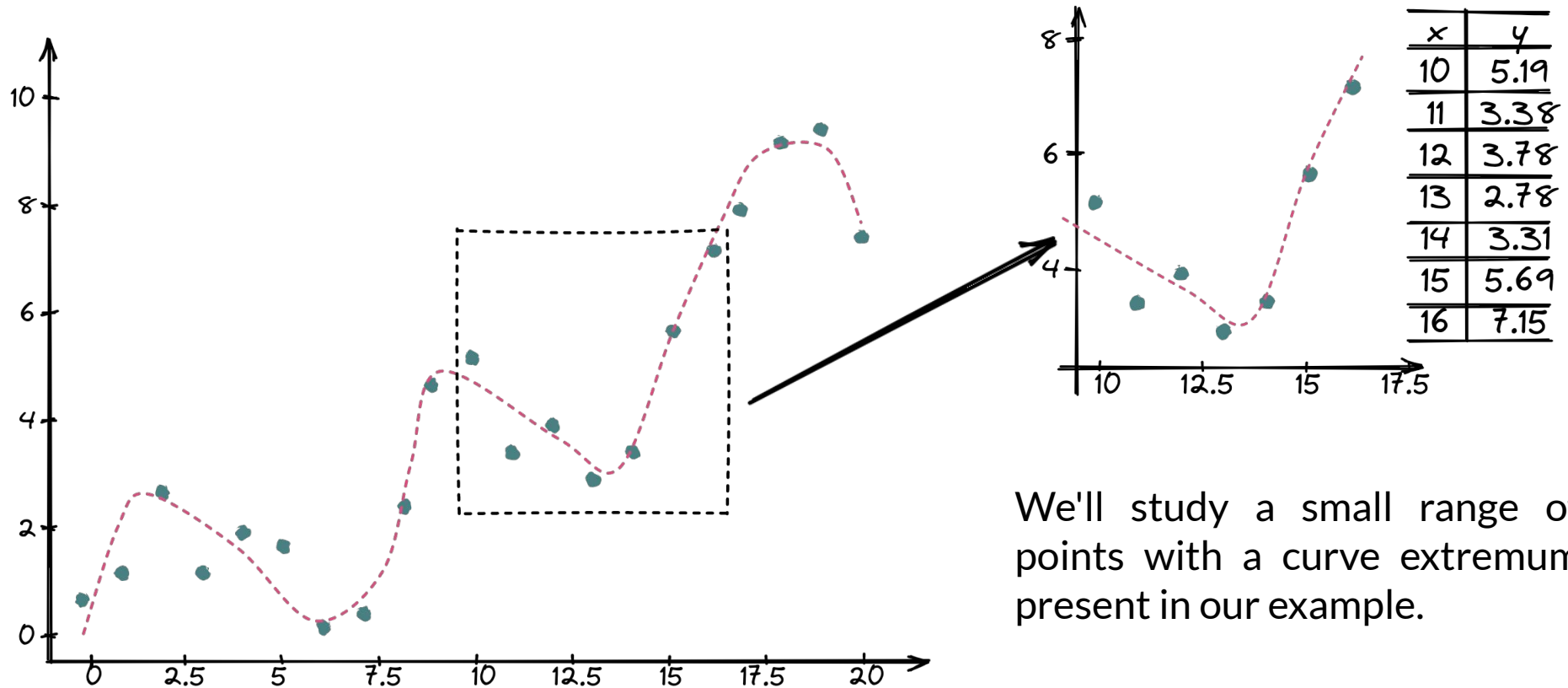
We'll study a small range of points with a curve extremum present in our example.

11

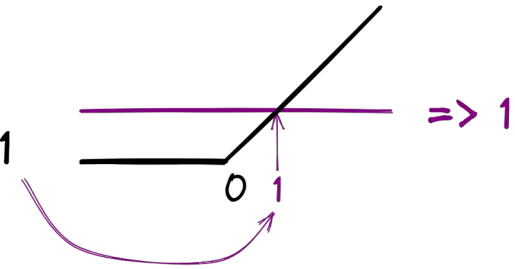# Case 1: Step 1a
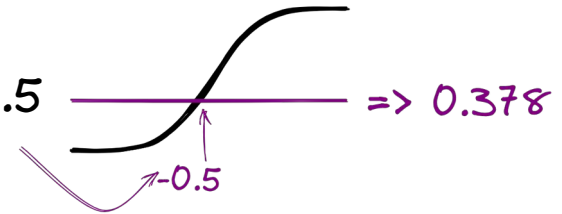

$i = 10$

$10 \times (-1) + 11 = 1$

$10 \times 0.2 - 2.5 = -0.5$

input   hidden   output

Case 1: Step 1b

input   hidden   output

$i \times (-1) + 11$

$i_{1,1} \times 1.1 + 1.9$

$i \times 0.2 + (-2.5)$

$i_{1,2} \times 2 + 1.4$

$i = 10$

$10 \times (-1) + 11 = 1$ $\Rightarrow 1$

$10 \times 0.2 - 2.5 = -0.5$ $\Rightarrow 0.378$

Case 1: Step 1c

$i \times (-1) + 11$

$i_{1,1} \times 1.1 + 1.9$

$i \times 0.2 + (-2.5)$

$i_{1,2} \times 2 + 1.4$

input   hidden   output

$i = 10$

$1 \times 1.1 + 1.9 = 3$

$0.378 \times 2 + 1.4 = 2.156$

Case 1: Step 1d

$i = 10$

$i \times (-1) + 11$

$i_{1,1} \times 1.1 + 1.9$

$i \times 0.2 + (-2.5)$

$i_{1,2} \times 2 + 1.4$

input   hidden   output

$\Rightarrow 5.061$

$0 \quad 5.056$

$sum(3, \ 2.156) - 0.1 = 5.056$
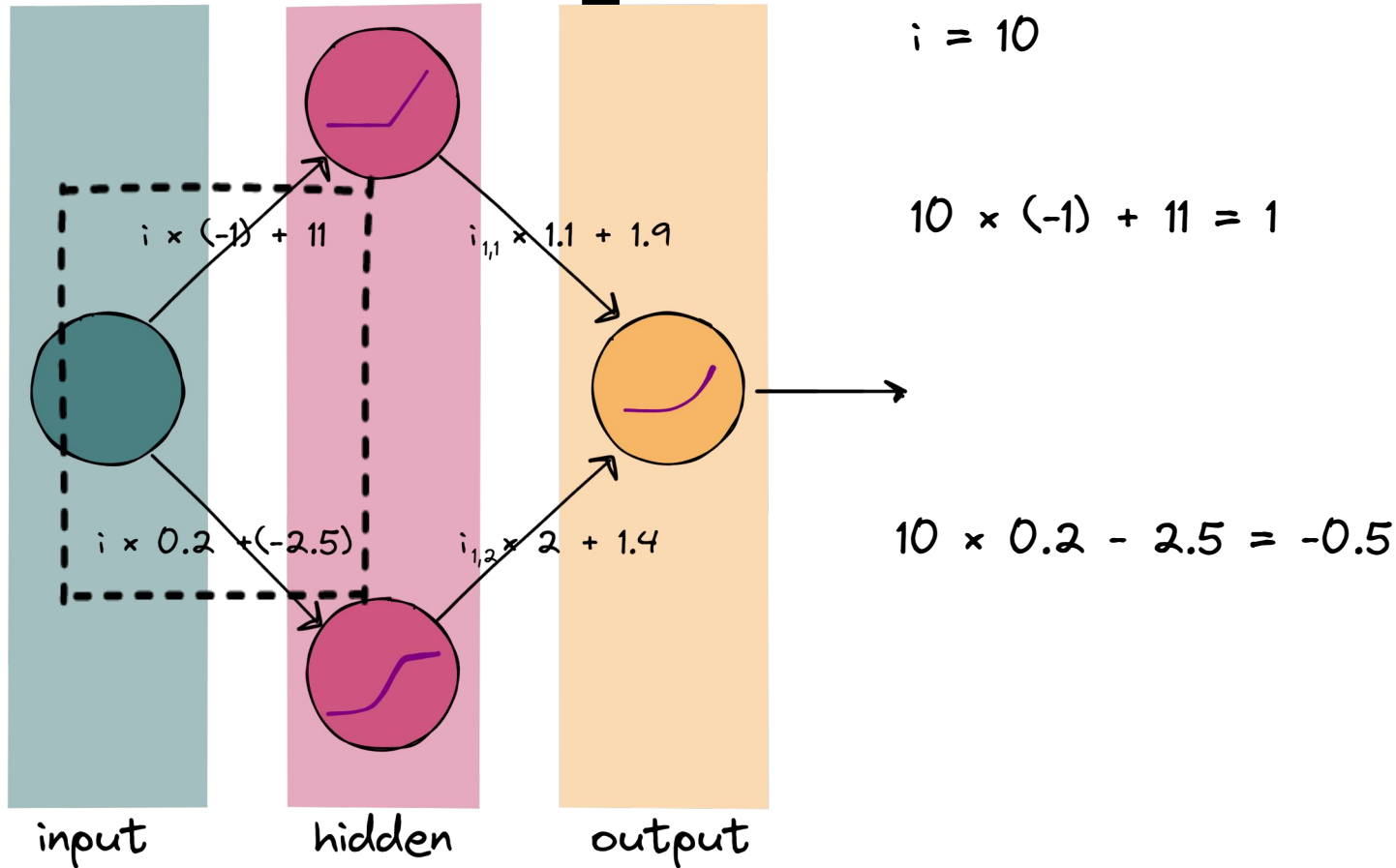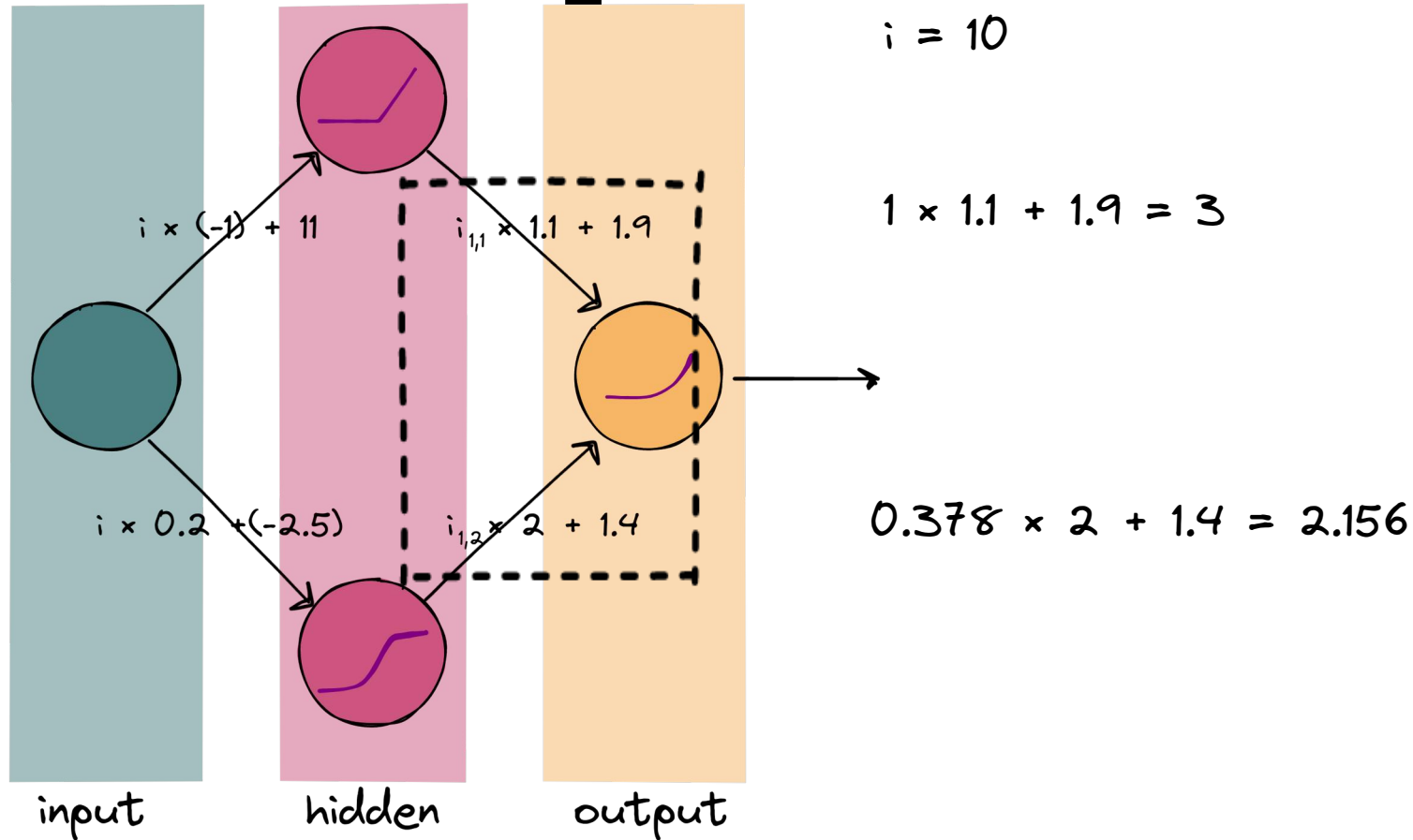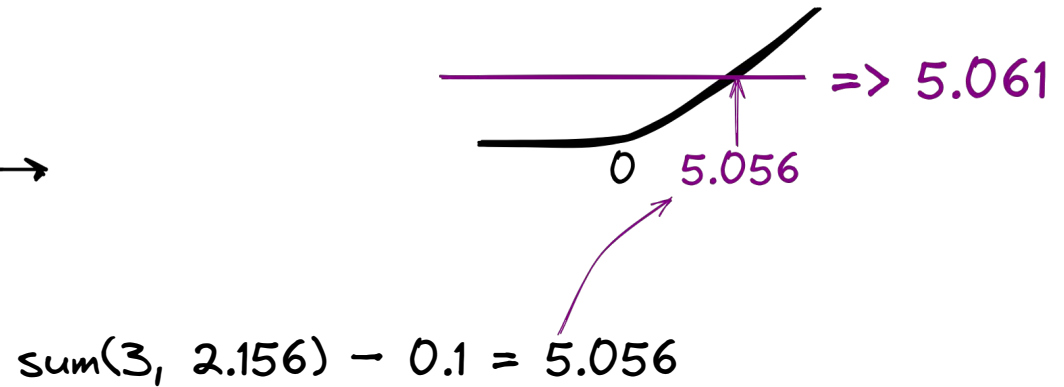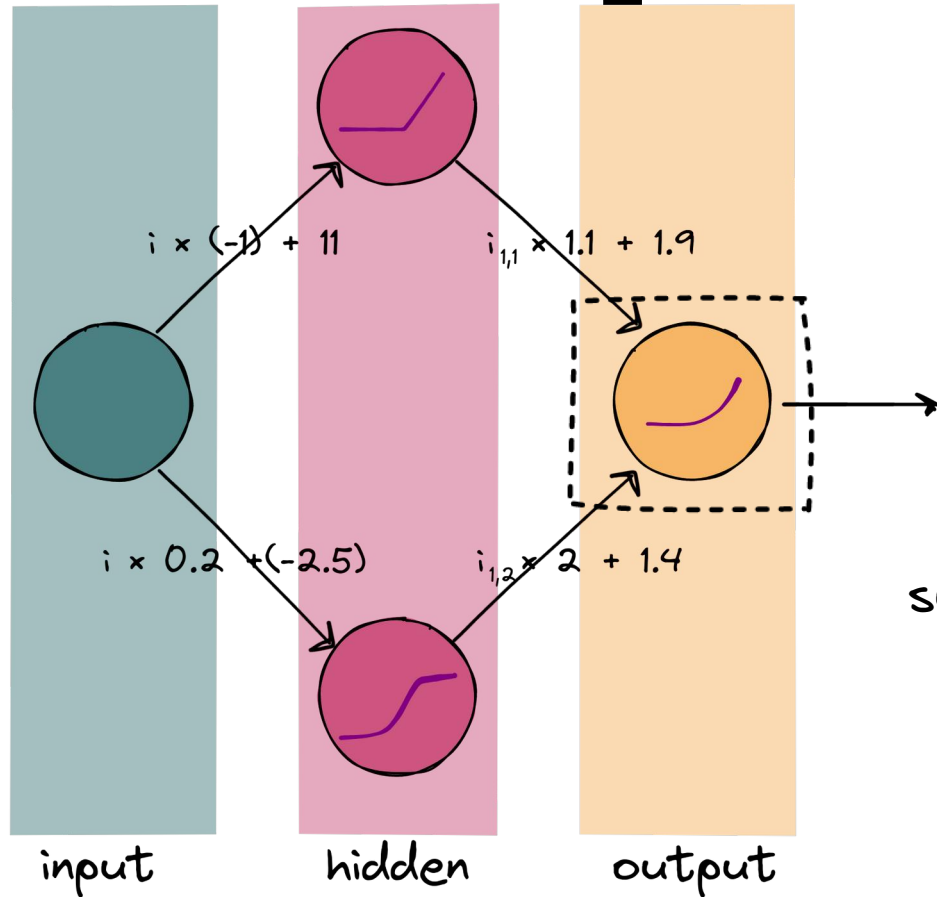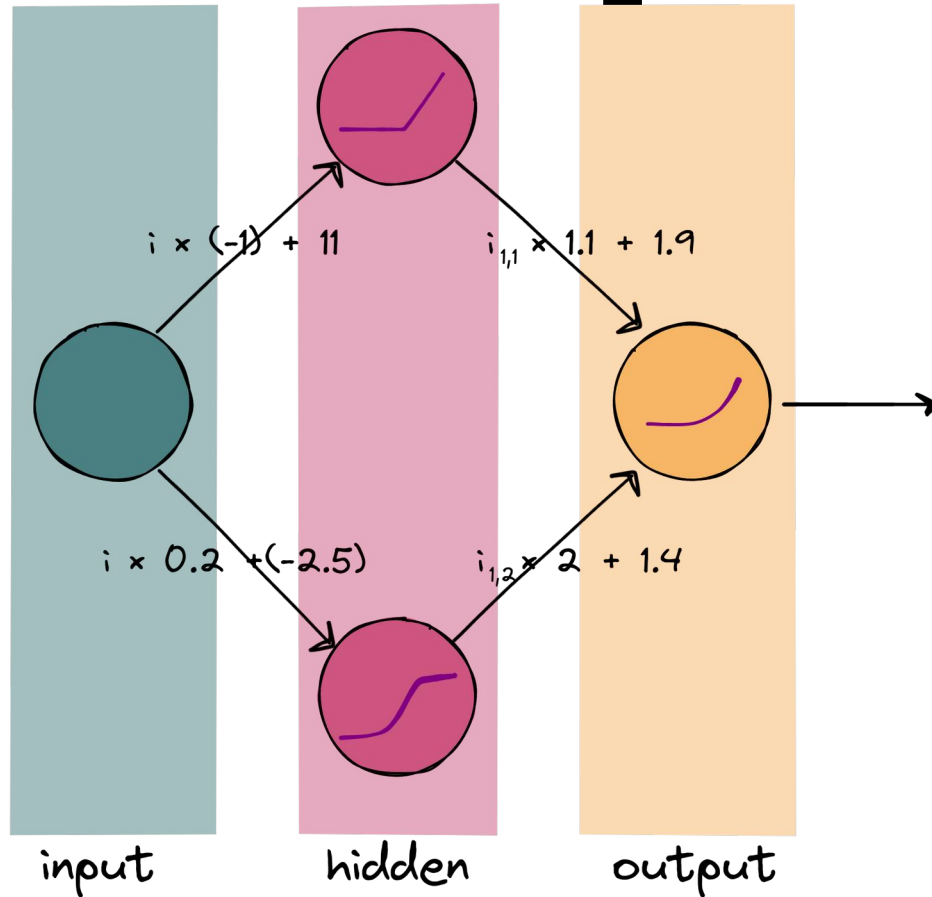
Case 1: Step 1e

x	y_true	y_pred
10	5.19	5.061
11	3.38	
12	3.78	
13	2.78	
14	3.31	
15	5.69	
16	7.15	
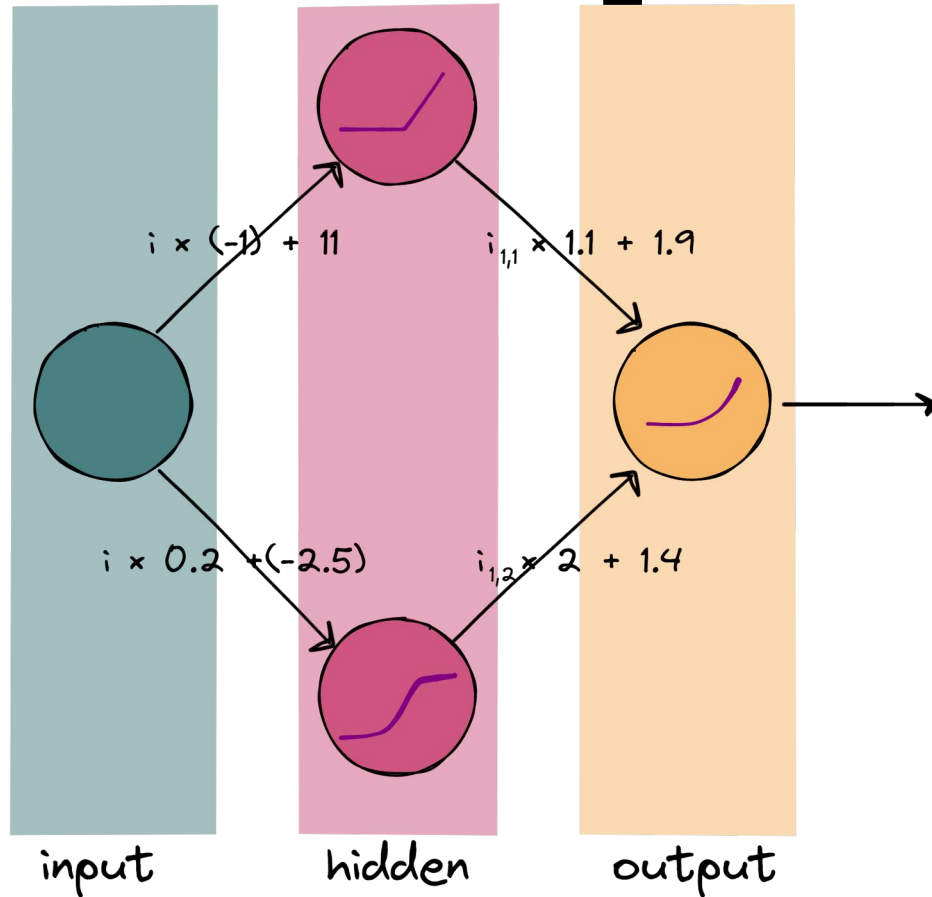
input   hidden   output

i × (-1) + 11 $i_{1,1}$ × 1.1 + 1.9

i × 0.2 +(-2.5) $i_{1,2}$ × 2 + 1.4

16

Case 1: Step 7e

x	y_true	y_pred
10	5.19	5.061
11	3.38	4.068
12	3.78	4.166
13	2.78	4.264
14	3.31	4.362
15	5.69	4.457
16	7.15	4.547

input   hidden   output

i × (-1) + 11 $i_{1,1}$ × 1.1 + 1.9

i × 0.2 + (-2.5) $i_{1,2}$ × 2 + 1.4

Case 1: explanation

x	y_true	y_pred
10	5.19	5.061
11	3.38	4.068
12	3.78	4.166
13	2.78	4.264
14	3.31	4.362
15	5.69	4.457
16	7.15	4.547

$i \times (-1) + 11$

$i_{1,1} \times 1.1 + 1.9$

$i \times 0.2 + (-2.5)$

$i_{1,2} \times 2 + 1.4$

input   hidden   output

Case 1: epoch 5 of 20

input   hidden   output

$i \times (-1) + 11$

$i_{1,1} \times 1.1 + 1.9$

$i \times 0.2 + (-2.5)$

$i_{1,2} \times 2 + 1.4$

19

Case 1: epoch 20 of 20

input　　　hidden　　　output

$i \times w1 + b1$

$i_{1,1} \times w3 + b3$

$i \times w2 + b2$

$i_{1,2} \times w4 + b4$