DASIL Python Workshop # 3

October 11th 2022

Introduction to Machine Learning Part #3

Yusen He, PhD - DASIL Data Scientist

Prof. Julia Bauder – Director of DASIL

Martin Pollack – DASIL Post Bachelor Fellow

DATA ANALYSIS & SOCIAL INQUIRY LAB
GRINNELL COLLEGE

Intro to Machine Learning Part #3

AGENDA

Introduction to Artificial Neural Network

Parameters & Functions in ANN

How Neural Network Learn

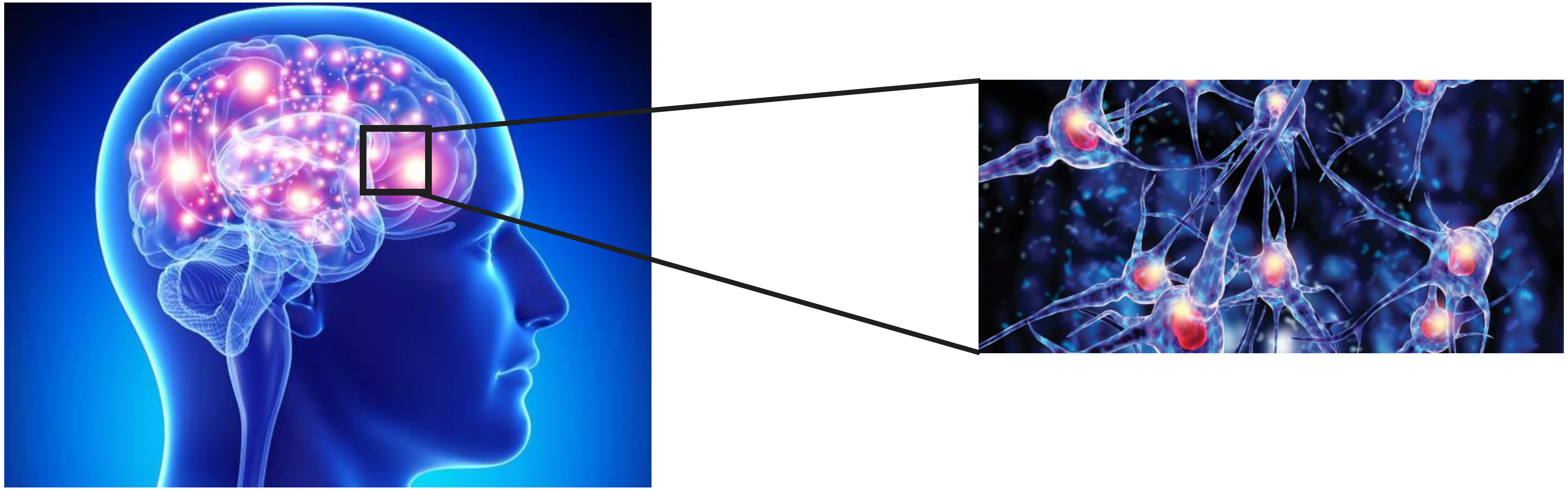
Other Types of Neural Networks

Model Evaluation

DATA ANALYSIS & SOCIAL INQUIRY LAB
GRINNELL COLLEGE

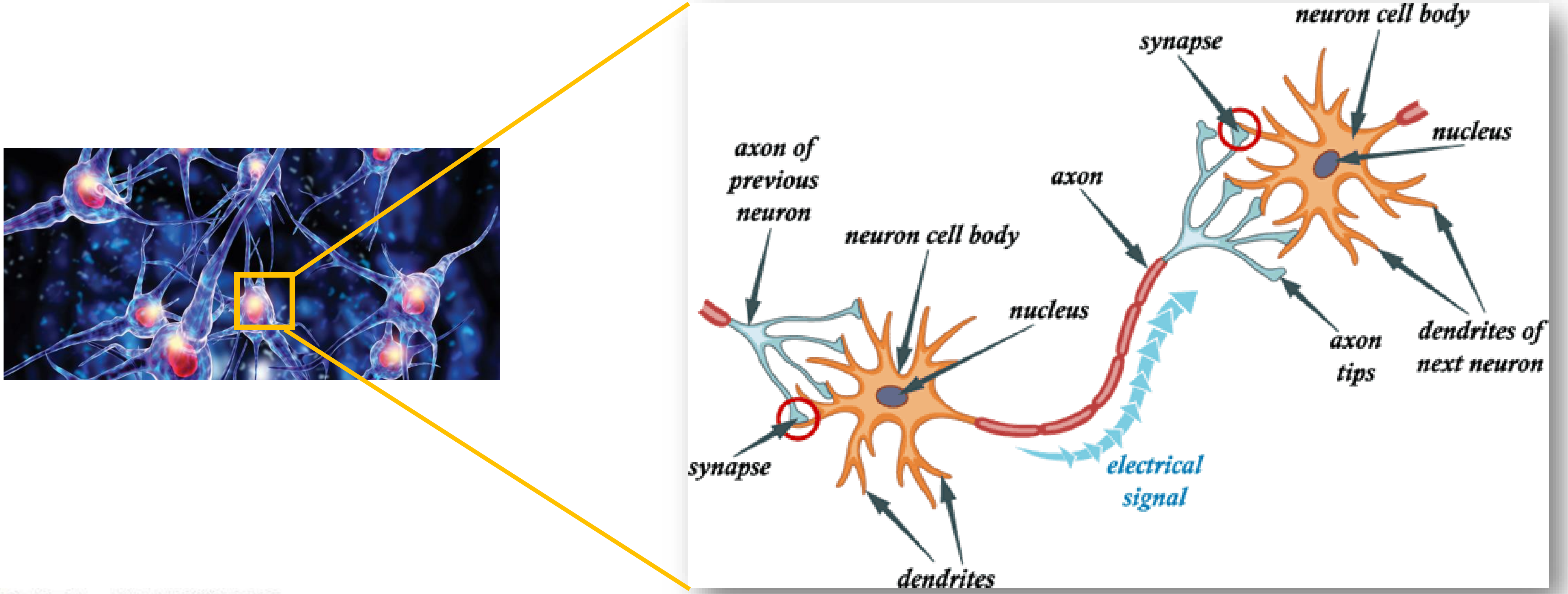
Introduction to Artificial Neural Network

- How human brain learn knowledge?



Introduction to Artificial Neural Network

- How human brain learn knowledge?



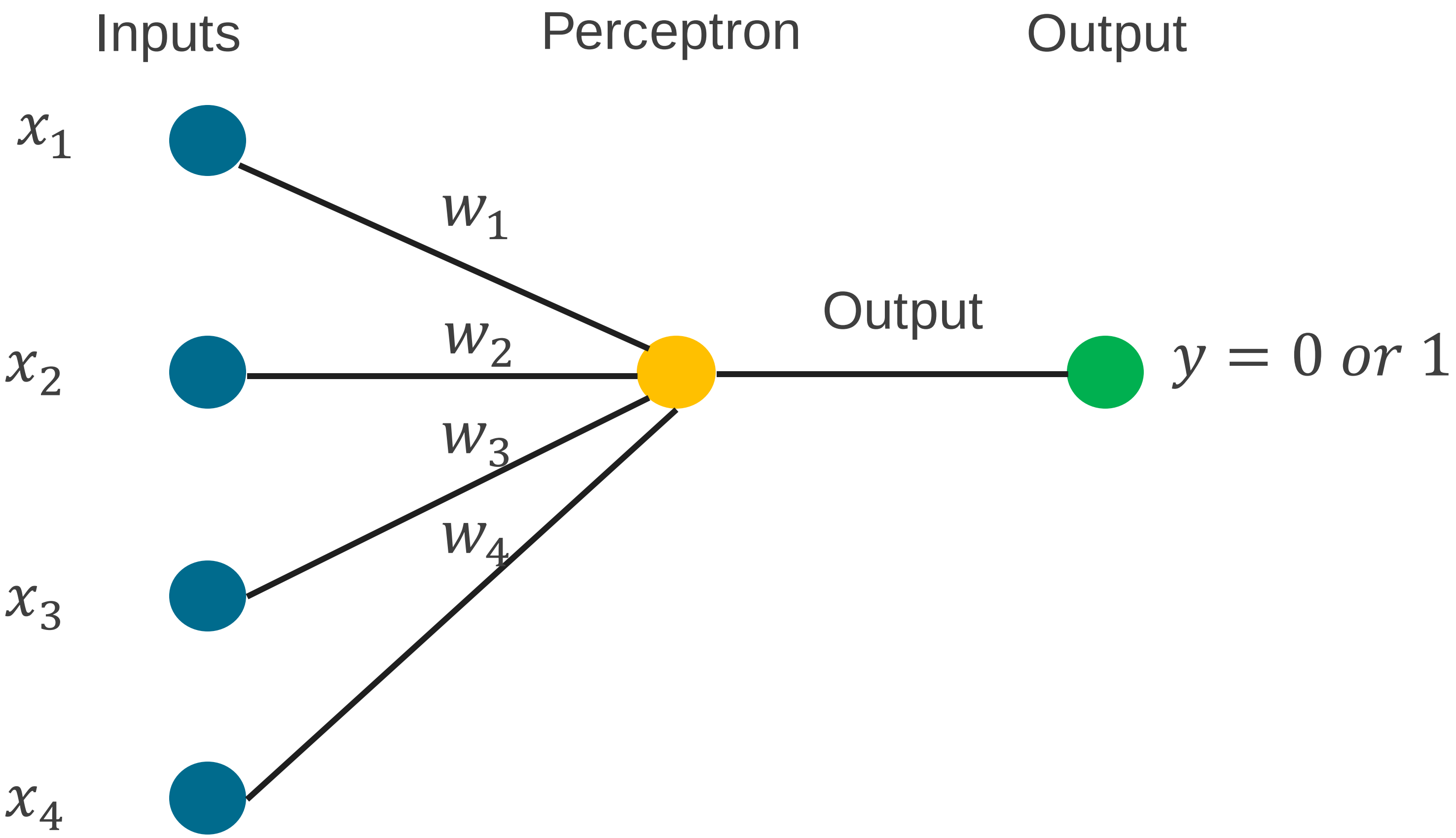
Neural Network Structure

Introduction to Artificial Neural Network

First Perceptron by Frank Rosenblatt (1957)

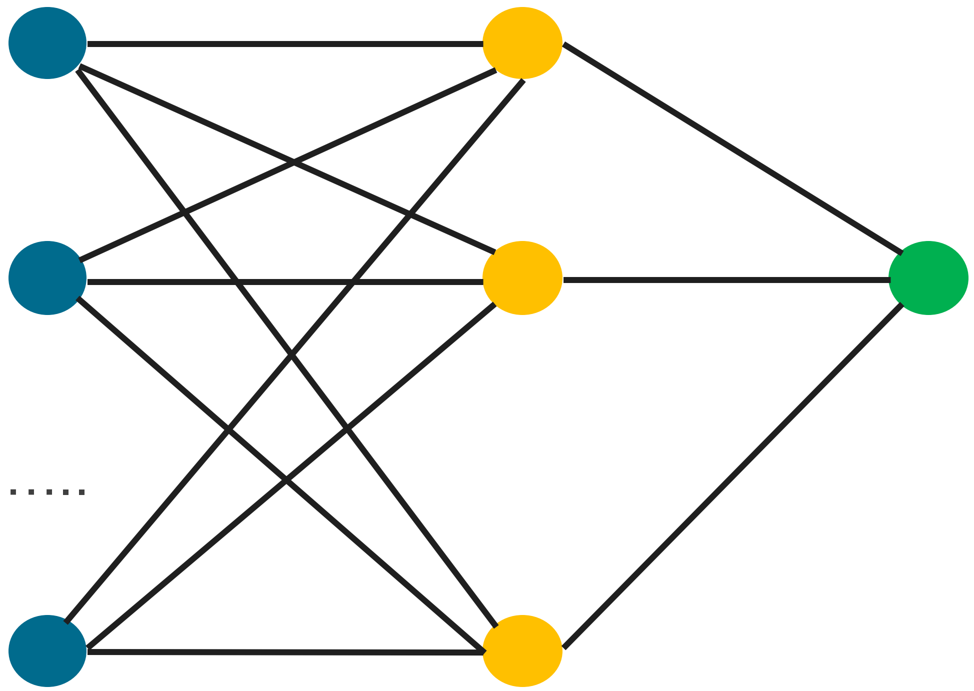
Introduction to Artificial Neural Network

A Perceptron Structure



Introduction to Artificial Neural Network

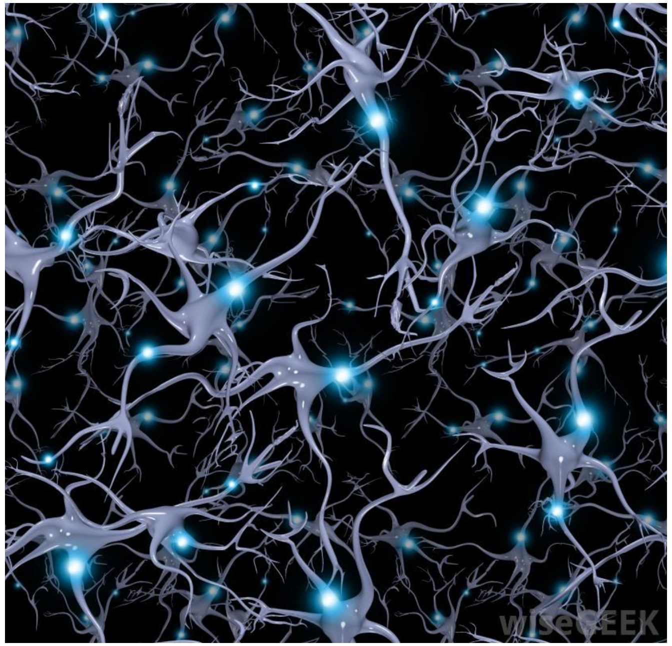
Minsky & Papert (1969): First Multilayer Perceptron (ANN)



DATA ANALYSIS & SOCIAL INQUIRY LAB
GRINNELL COLLEGE

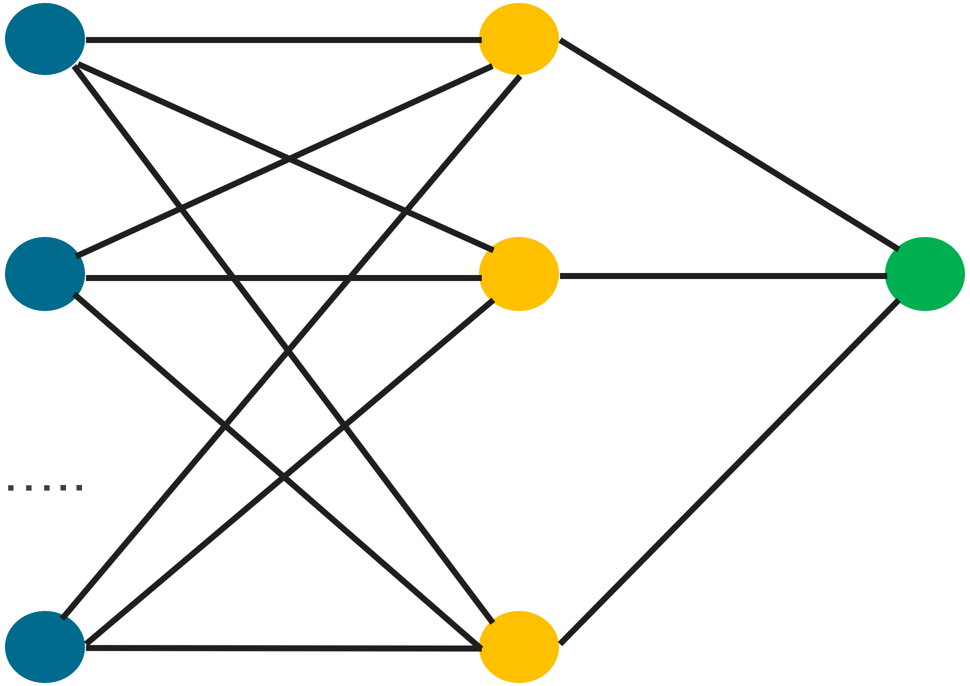
Introduction to Artificial Neural Network

Discussion: What is the difference?



Human neural network

Versus



ANN

DATA ANALYSIS & SOCIAL INQUIRY LAB GRINNELL COLLEGE

Intro to Machine Learning Part #3

AGENDA

Introduction to Artificial Neural Network

Parameters & Functions in ANN

How Neural Network Learn

Other Types of Neural Networks

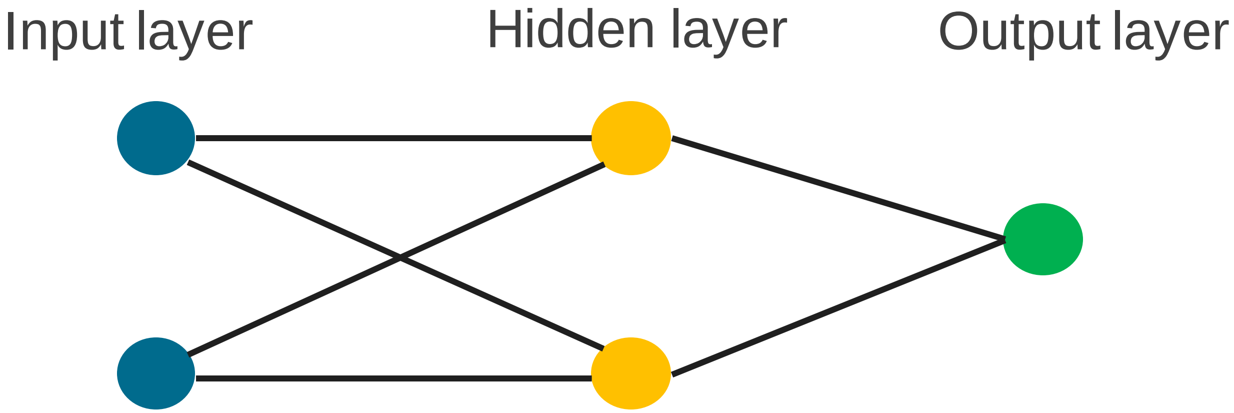
Model Evaluation

DATA ANALYSIS & SOCIAL INQUIRY LAB
GRINNELL COLLEGE

Parameters & Functions in ANN

Three layers:

- Input layer
- Hidden layer
- Output layer

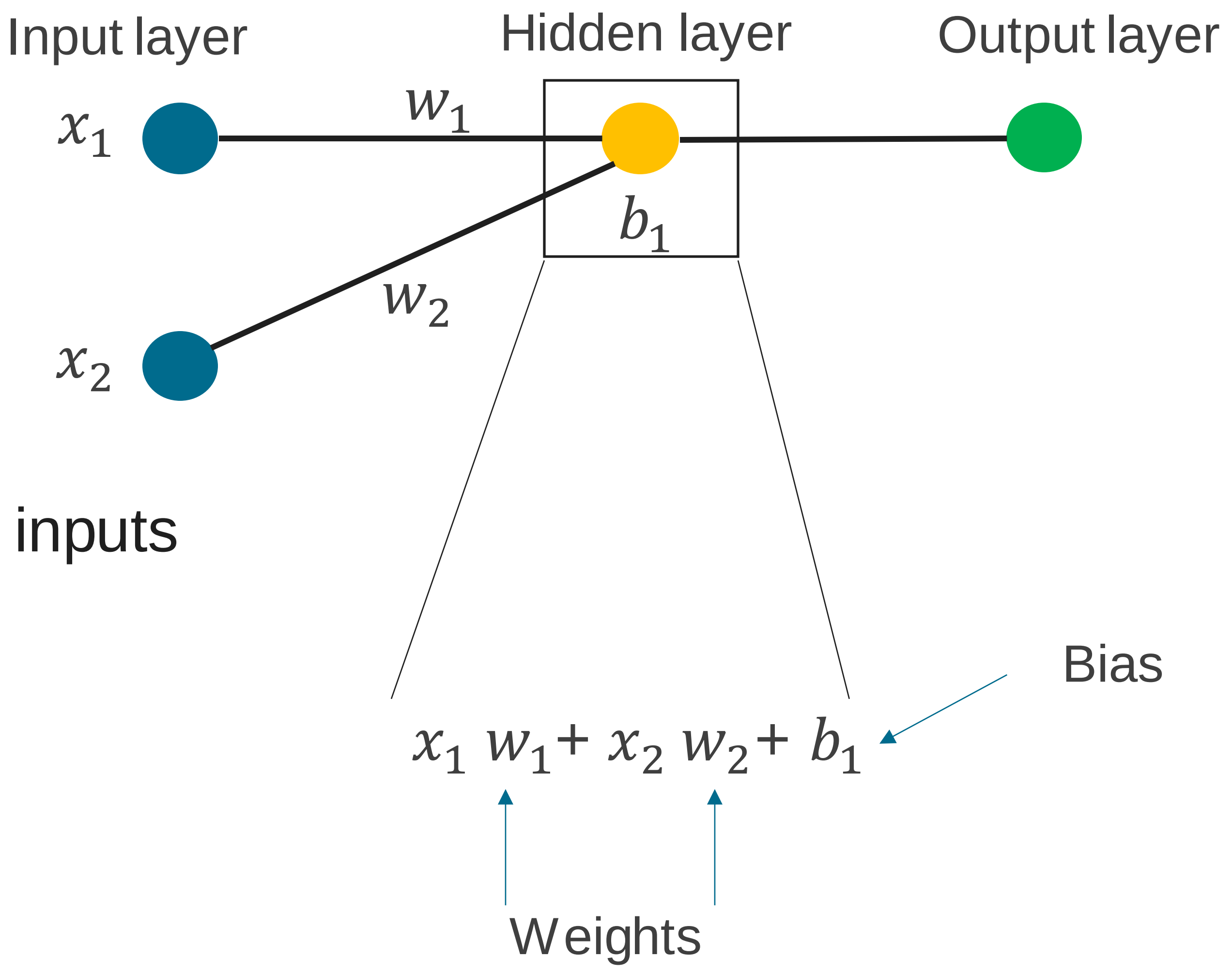


Parameters & Functions in ANN

Weights & Bias

Weights: Converts input data

Bias: Add on sum of weights x inputs



Input layer

Hidden layer

Output layer

x_1 w_1

x_2 w_2

b_1

$$x_1\,w_1 + x_2\,w_2 + b_1$$

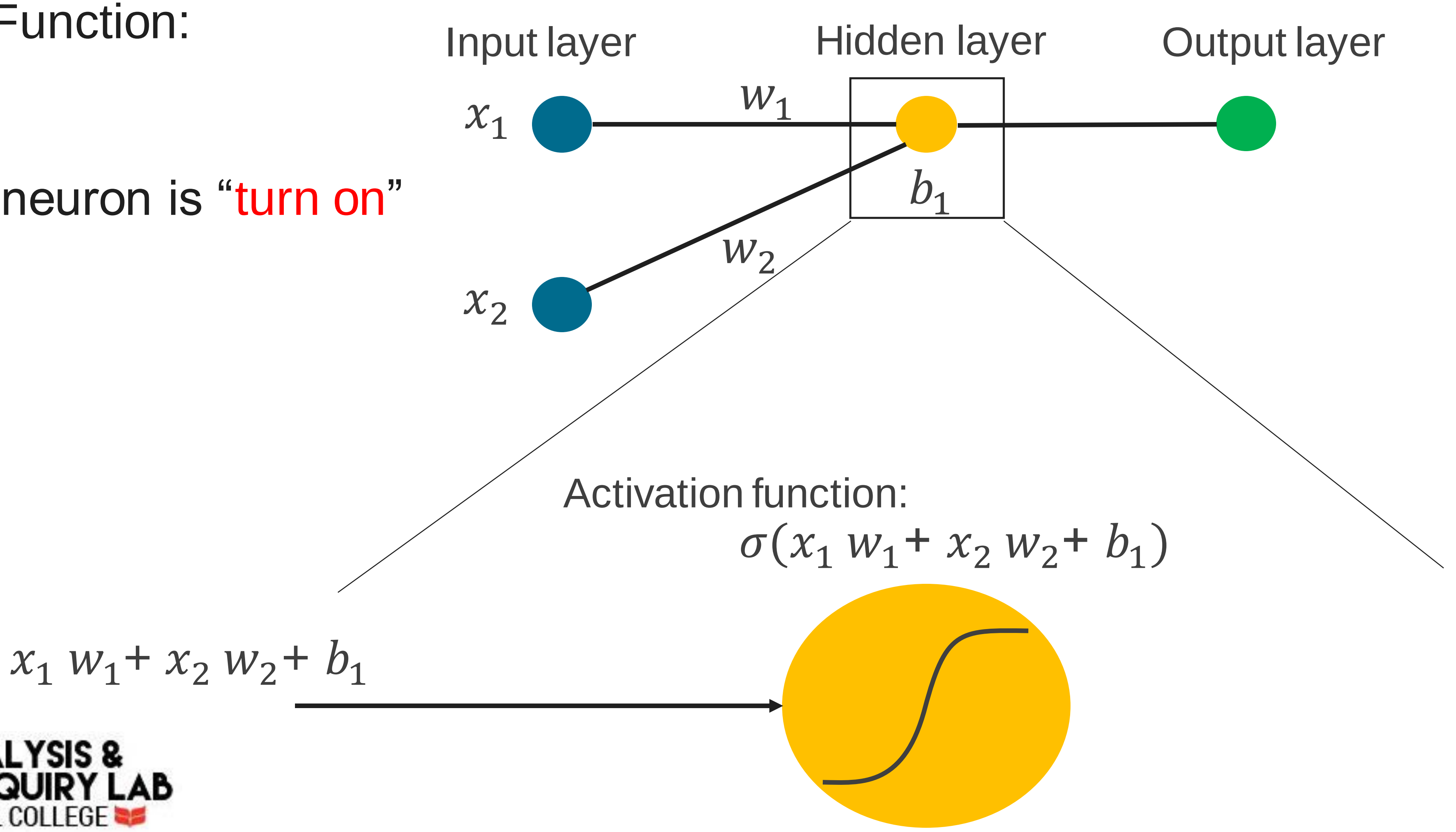
Bias

Weights

Parameters & Functions in ANN

Activation Function:

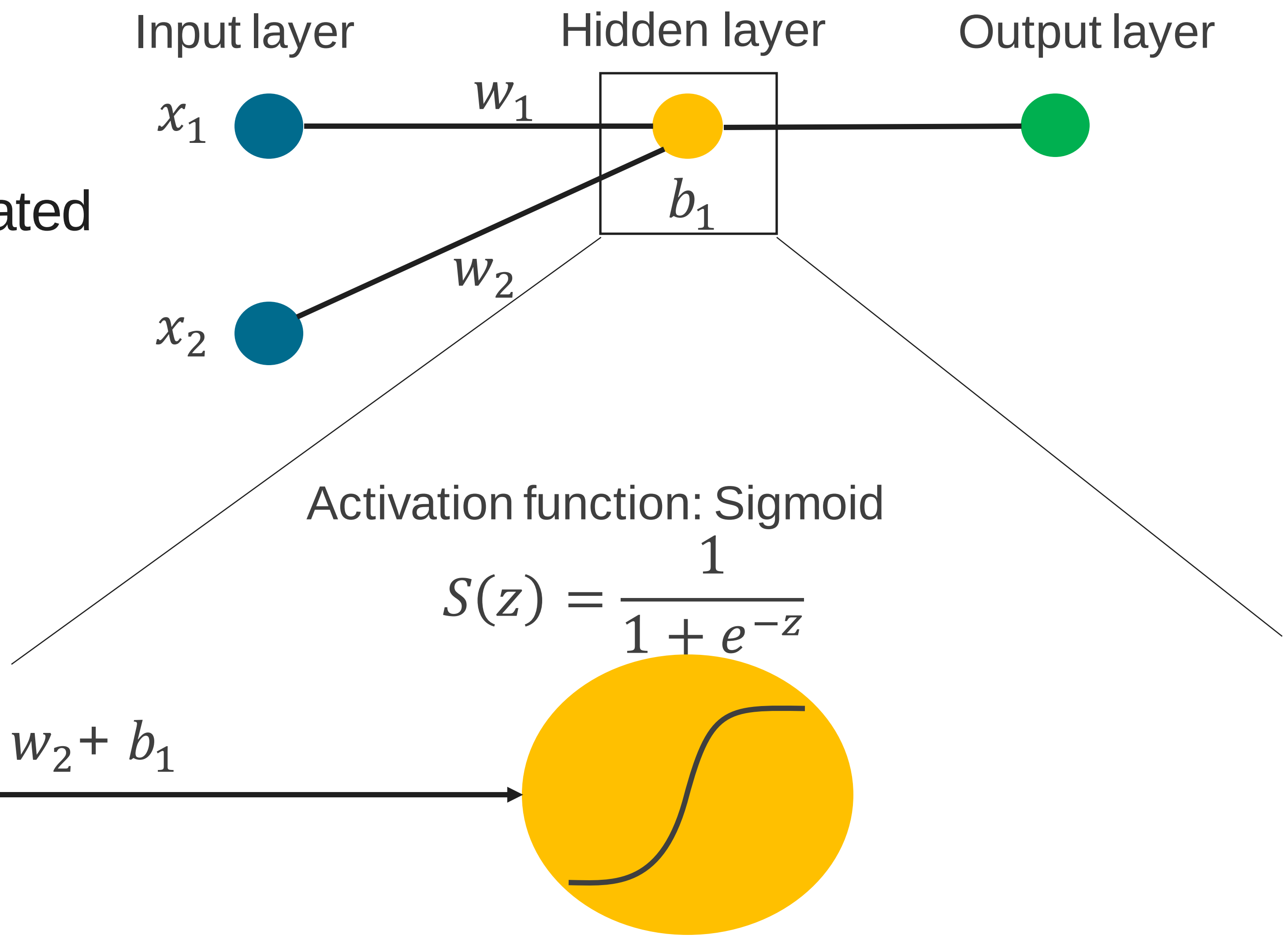
Whether a neuron is “turn on”



Parameters & Functions in ANN

Activation Function:

Whether a neuron is activated



Activation function: Sigmoid

$$S(z) = \frac{1}{1 + e^{-z}}$$

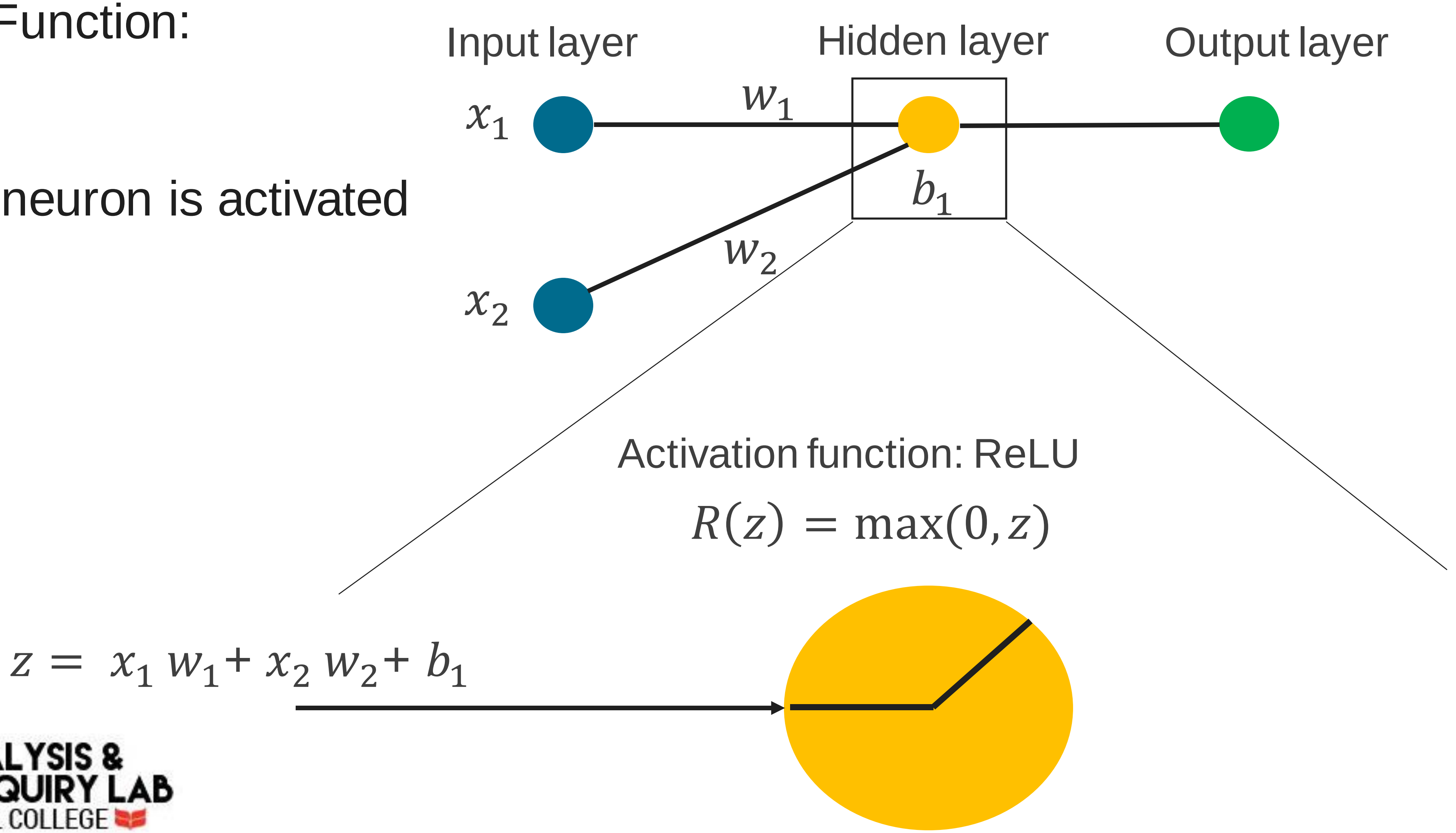
$$z = x_1 w_1 + x_2 w_2 + b_1$$

DATA ANALYSIS & SOCIAL INQUIRY LAB
GRINNELL COLLEGE

Parameters & Functions in ANN

Activation Function:

Whether a neuron is activated



Activation function: ReLU

$$R(z) = \max(0, z)$$

$$z = x_1 w_1 + x_2 w_2 + b_1$$

Parameters & Functions in ANN

In class exercise:

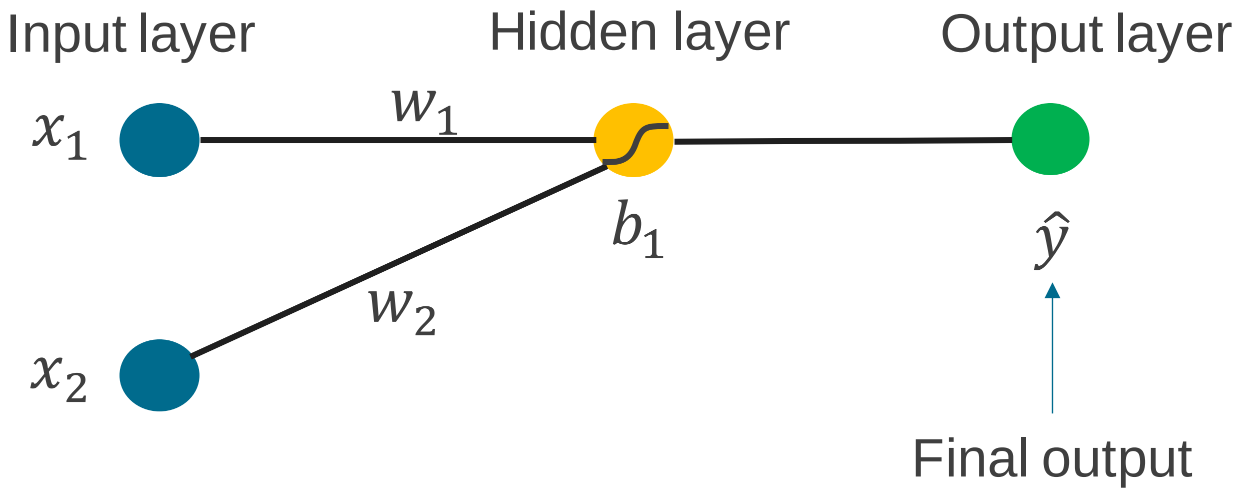
Suppose we have $z = x_1 w_1 + x_2 w_2 + b_1 = 0.4$

Should the neuron “turn on” if $\sigma(z) = \dfrac{1}{1 + e^{-z}}$

Should the neuron “turn on” if $\sigma(z) = \max(0, z)$

Parameters & Functions in ANN

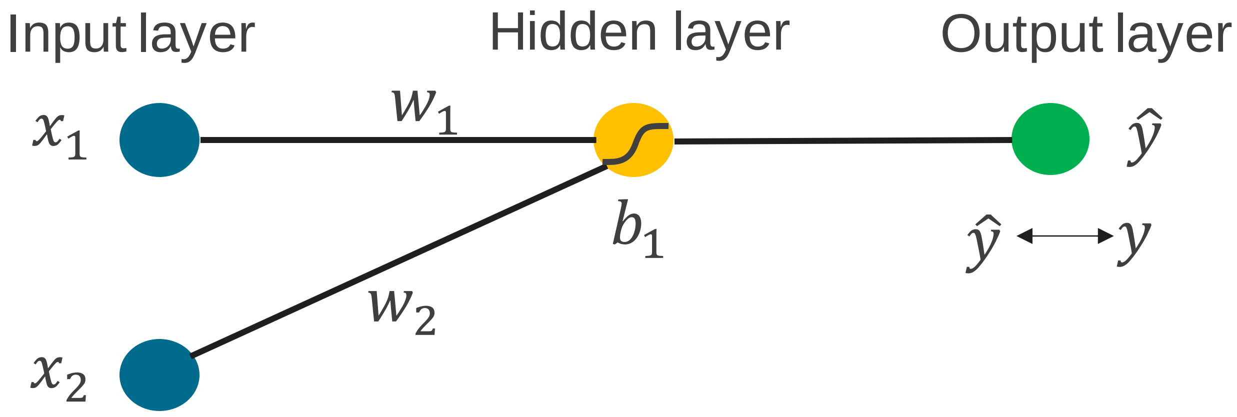
Output: The final prediction



Parameters & Functions in ANN

Loss Function:

How far the prediction is from truth



$Loss(\hat{y}, y)$

Parameters & Functions in ANN

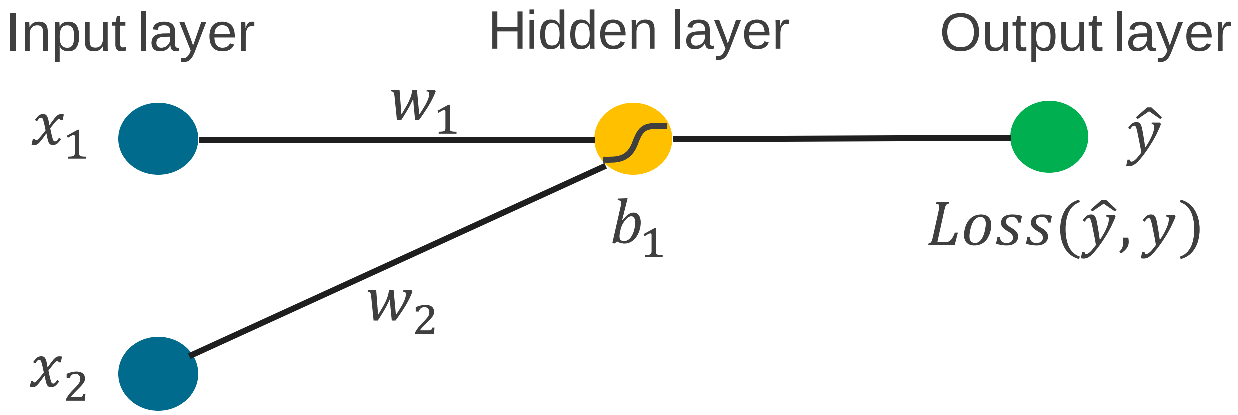
Regression Tasks:

Actual

Prediction

Mean Square Error (MSE) Loss: $$Loss(\hat{y}, y) = \frac{1}{N}(\hat{y} - y)^2$$

Sample size



Parameters & Functions in ANN

Classification Tasks:

Log-predicted (+) probability

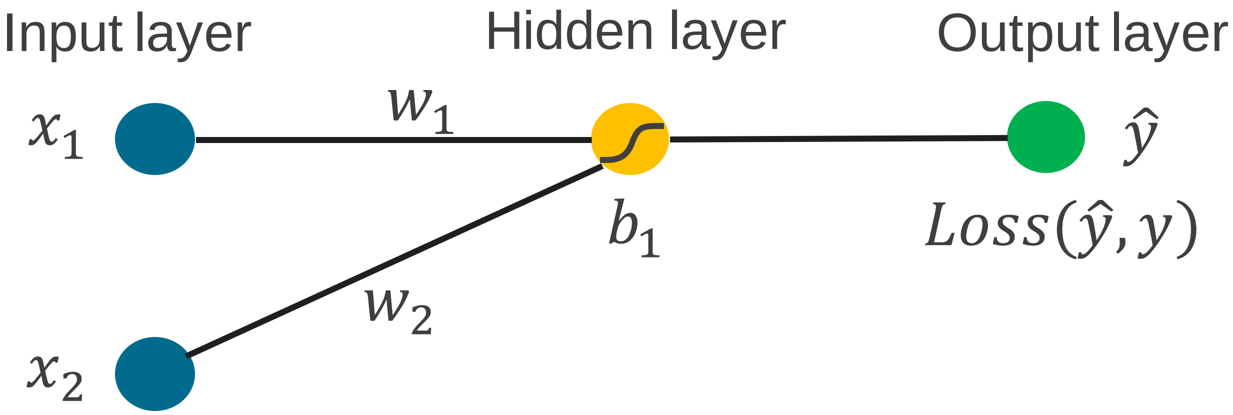
Log-predicted (-) probability

Actual positive

Actual negative

Binary Cross-Entropy (BCE) Loss: $Loss(\hat{y}, y) = \frac{1}{N}(y \times \log \hat{y} - (1 - y) \times \log(1 - \hat{y}))$

Sample size



Intro to Machine Learning Part #3

AGENDA

Introduction to Artificial Neural Network

Parameters & Functions in ANN

How Neural Network Learn

Other Types of Neural Networks

Model Evaluation

DATA ANALYSIS & SOCIAL INQUIRY LAB
GRINNELL COLLEGE

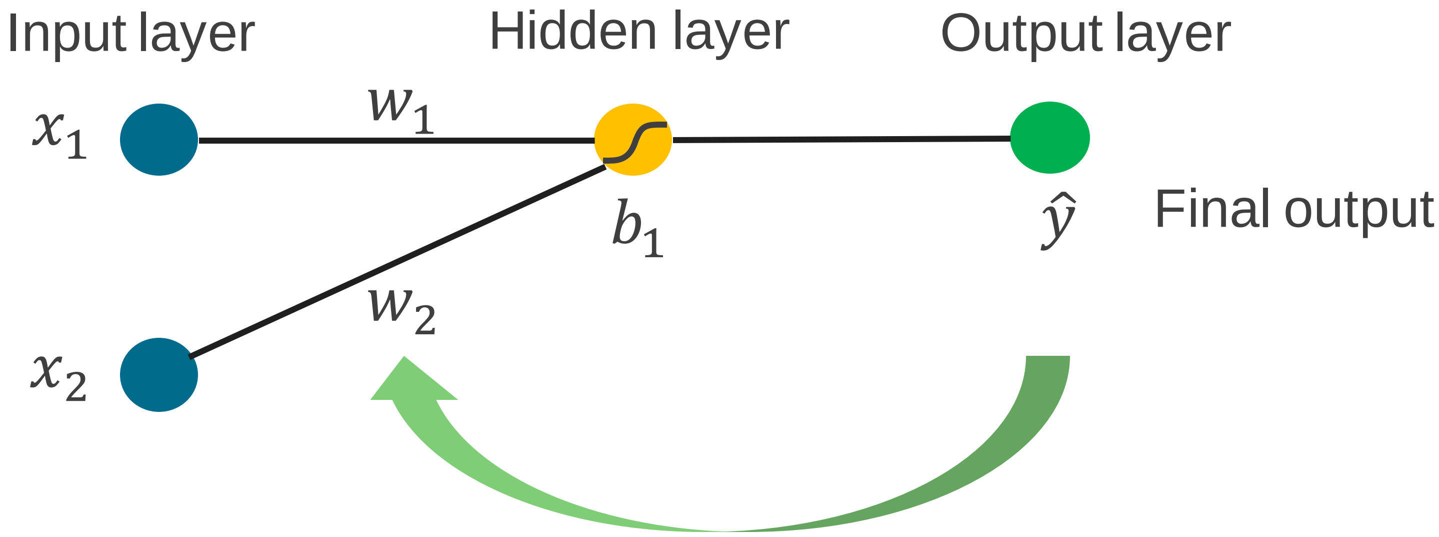
How Neural Network Learn?

Rumelhart, Hinton, and Williams. (1986).

- We need to learn the following ***parameters:***
  - Weights (w_i)
  - Biases(b_i)

- Start with random initial guesses

- Then do ***backpropagation***



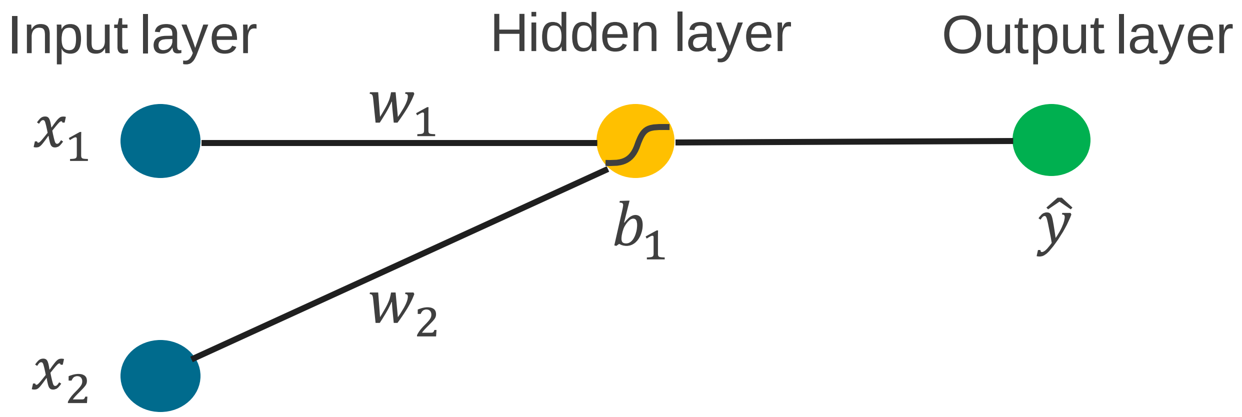
Differentiate, Multiply, Sum, Differentiate, Multiply, Sum, ….

DATA ANALYSIS & SOCIAL INQUIRY LAB GRINNELL COLLEGE

How Neural Network Learn?

Back Propagation:

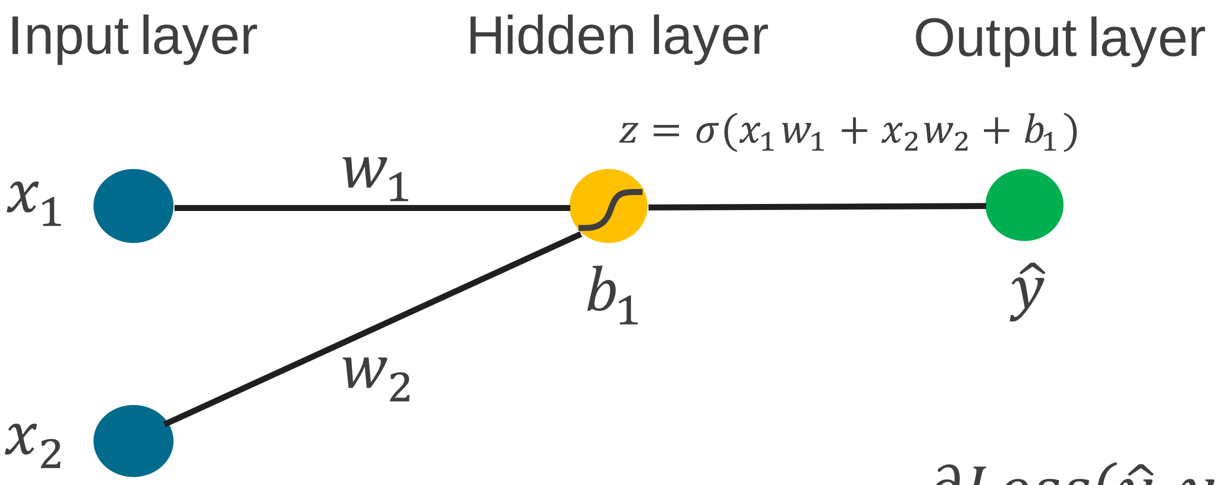
Step 1: Compute the loss function $Loss(\hat{y}, y)$



How Neural Network Learn?

Back Propagation:

Step 2: Compute the gradient of parameters (how much the loss function changes with respect to the parameters) via Chain Rule:



Input layer

Hidden layer

Output layer

$z = \sigma(x_1 w_1 + x_2 w_2 + b_1)$

x_1

w_1

b_1

$\hat{y}$

w_2

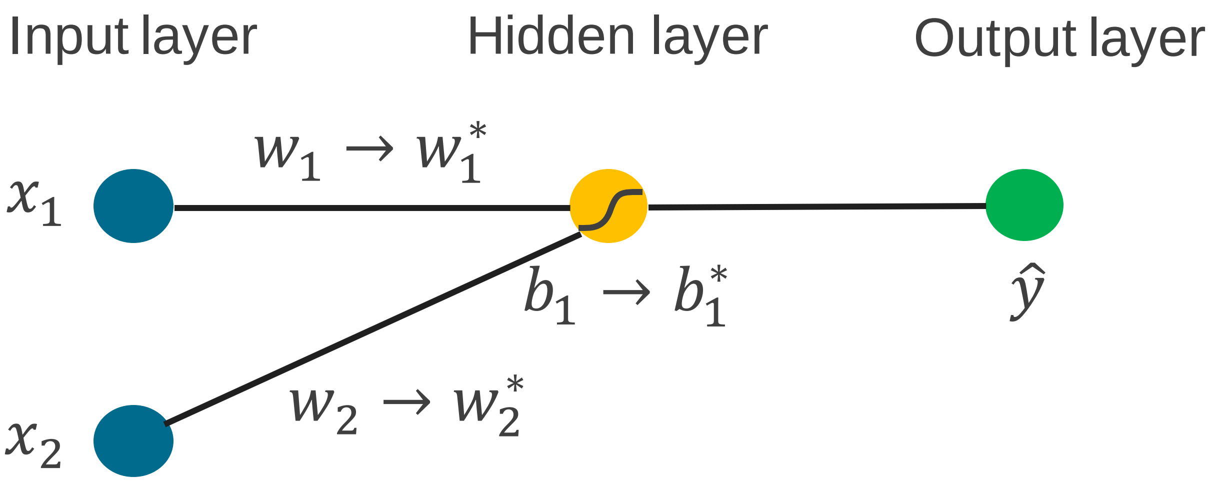
x_2

For example: $\nabla = \frac{\partial Loss(\hat{y}, y)}{\partial w_1} = \frac{\partial Loss(\hat{y}, y)}{\partial z} \times \frac{\partial z}{\partial w_1}$

How Neural Network Learn?

Back Propagation:

Step 3: Update the parameters using Gradient:



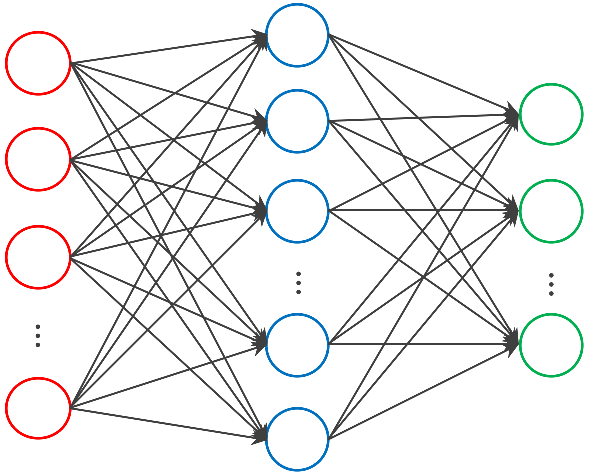
Learning rate

$$w_i^* = w_i - \alpha\nabla$$

...

Gradient

Q & A



DATA ANALYSIS & SOCIAL INQUIRY LAB GRINNELL COLLEGE