Fully Connected Neural Networks

Natalia Dubljevic (guest lecturer)
ENEL 645
Biomedical Engineering
Schulich School of Engineering

F2024

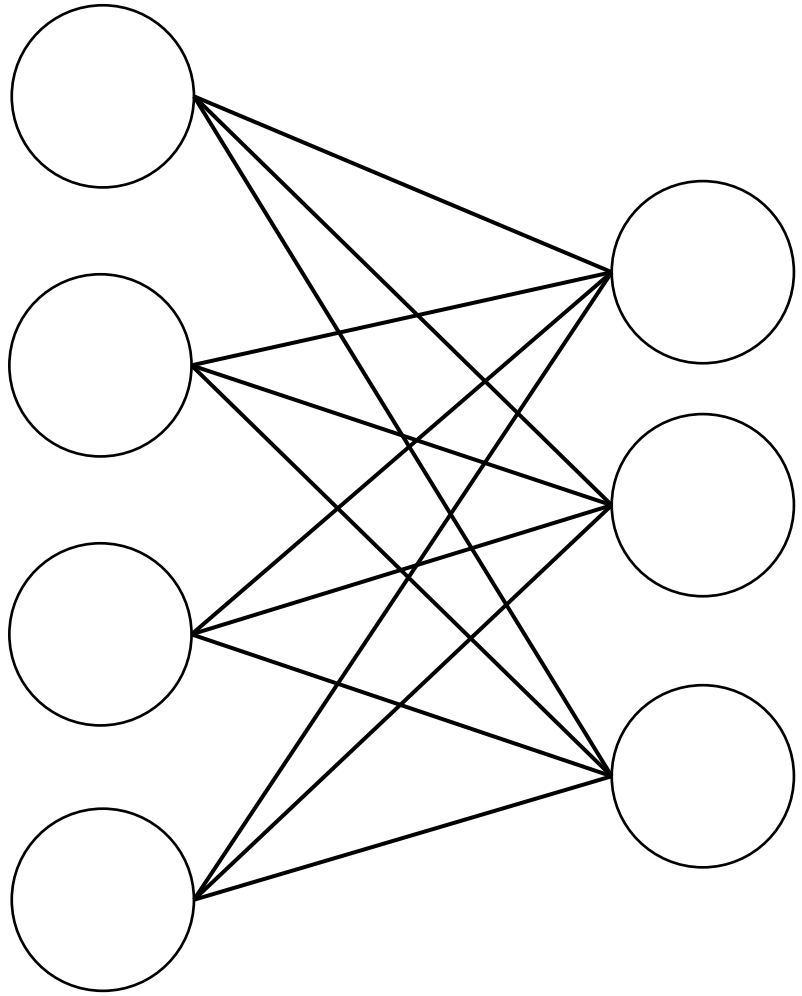
Outline

- Learning Goals
- Fully Connected Neural Networks
  - Single-layer Network
  - Multi-layer Network
- Summary
- Tutorial

Learning Goals

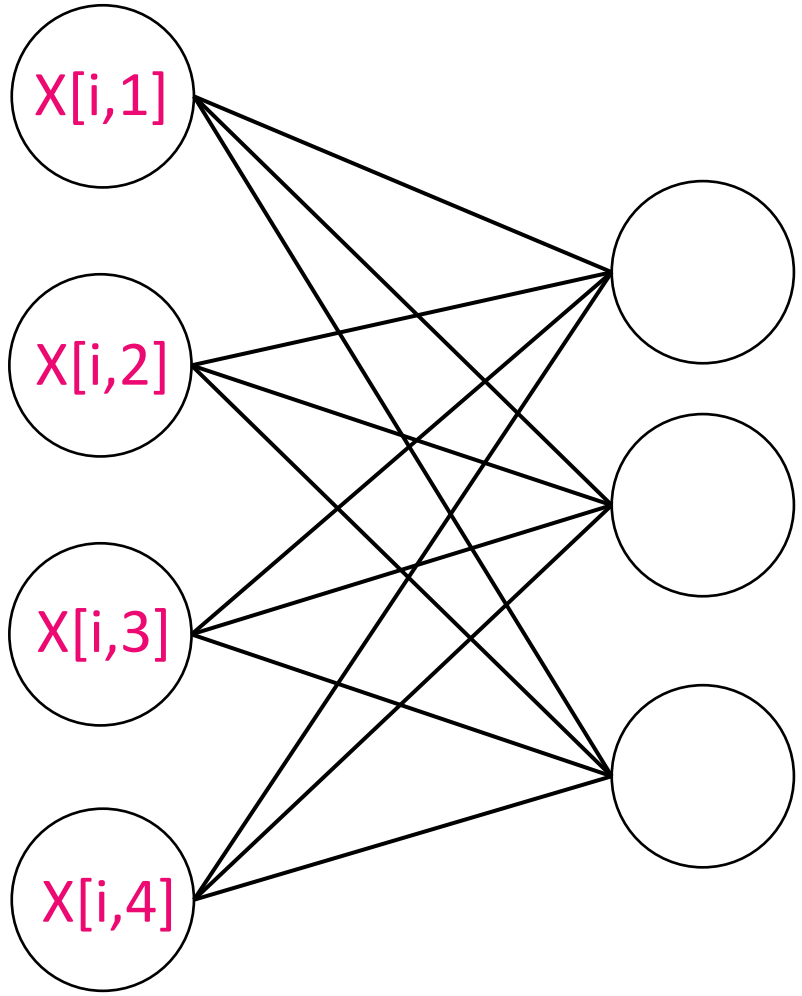
- Introduce fully connected neural networks
- Learn how to compute the number of parameters of your model

Fully connected neural networks (FCNN)



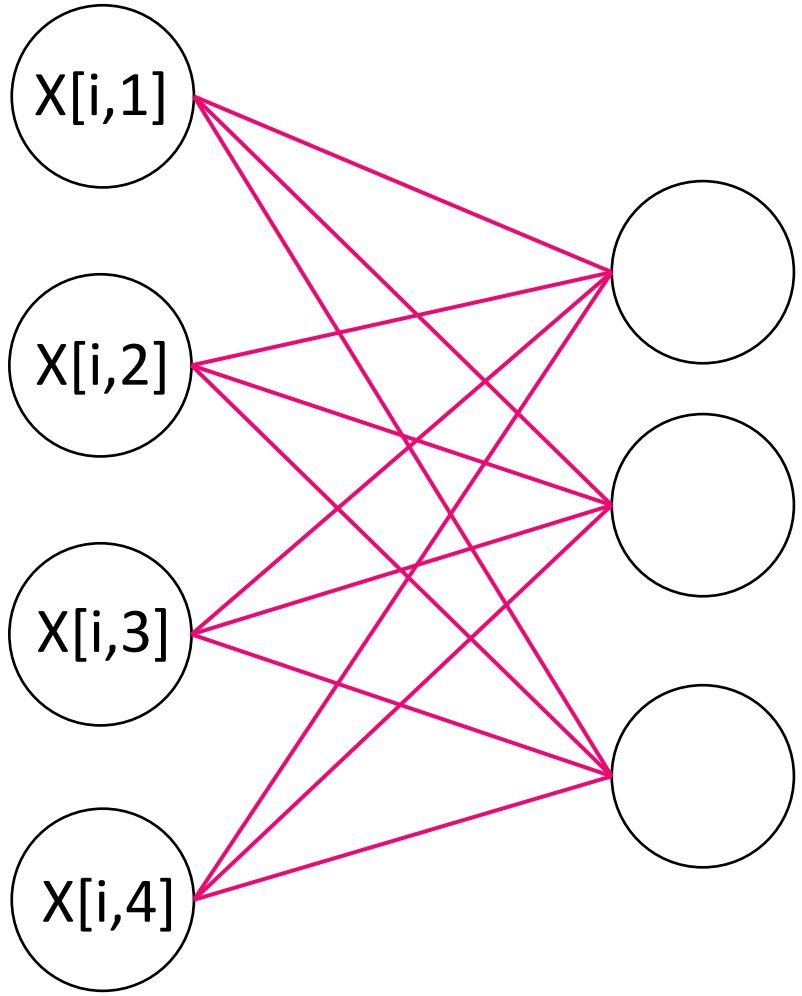
UNIVERSITY OF CALGARY

Inputs



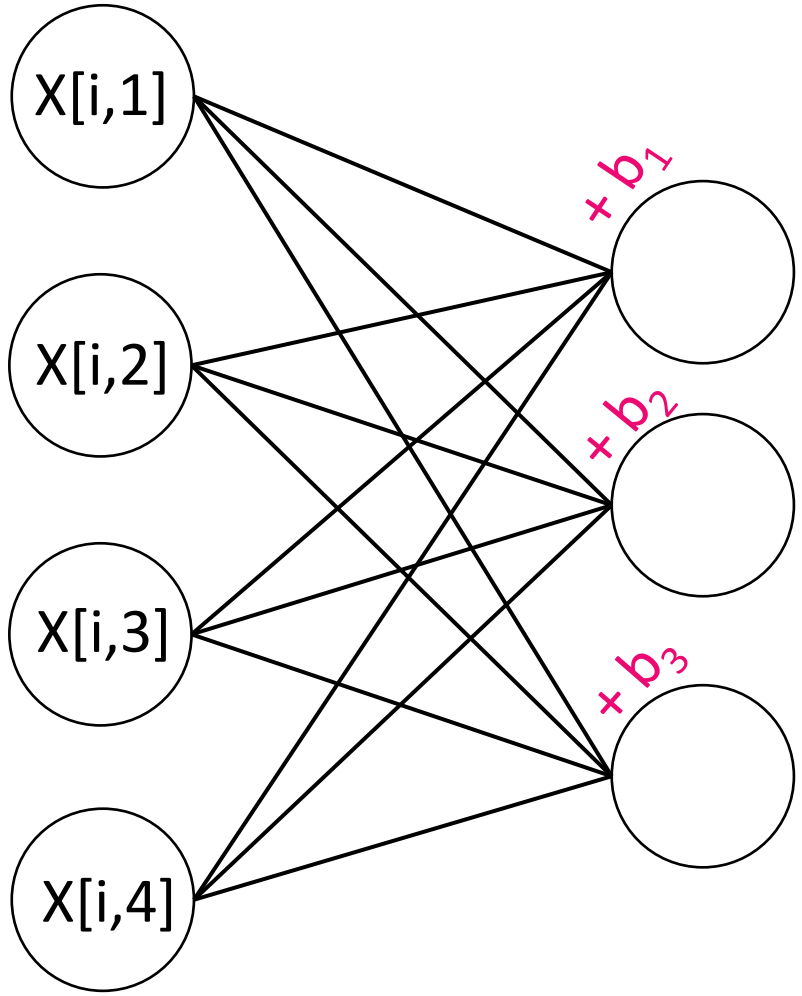
$$X = \underbrace{\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1M} \\ x_{21} & x_{22} & \cdots & x_{2M} \\ \cdots & \cdots & \cdots & \cdots \\ x_{N1} & x_{N2} & \cdots & x_{NM} \end{bmatrix}}_{\text{N samples with M features}}$$

Weights



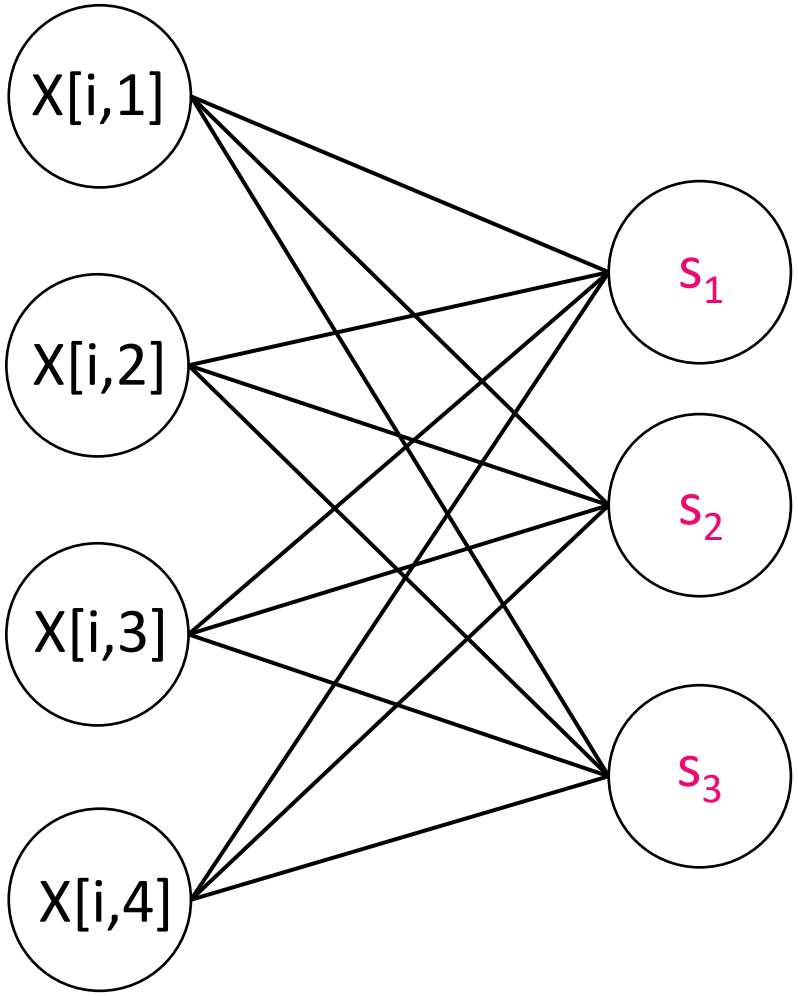
$$W = \underbrace{\begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1M} \\ w_{21} & w_{22} & \cdots & w_{2M} \\ \cdots & \cdots & \cdots & \cdots \\ w_{N1} & w_{N2} & \cdots & w_{NM} \end{bmatrix}}_{\text{N outputs M inputs}}$$

Bias



$$b = \underbrace{\begin{bmatrix} b_1 \\ \cdots \\ b_M \end{bmatrix}}_{\text{M outputs}}$$

The Neuron Model



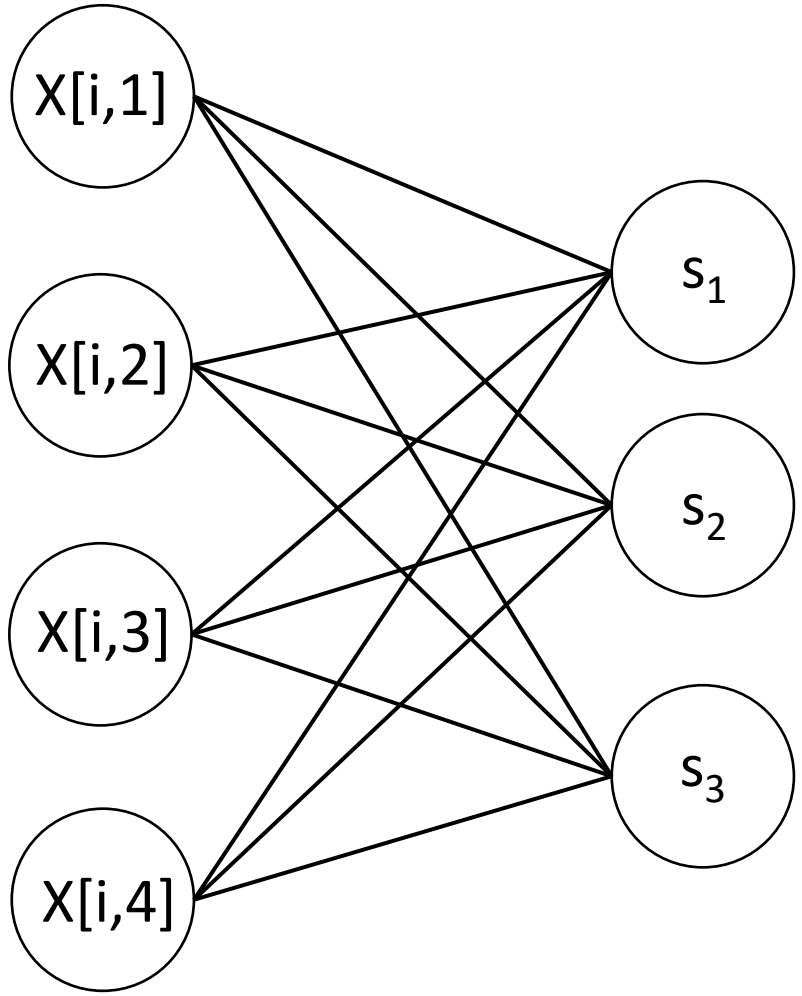
$$b_i + \sum_{j=1}^{M} w_{ij}\, X[i,j] \quad \sigma \quad \rightarrow s_i$$

Linear operation

Non-linear operation (activation)

UNIVERSITY OF CALGARY

Outputs



$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_N \end{bmatrix} \qquad \hat{Y} = \begin{bmatrix} \widehat{y_1} \\ \widehat{y_2} \\ \dots \\ \widehat{y_N} \end{bmatrix}$$

True Labels

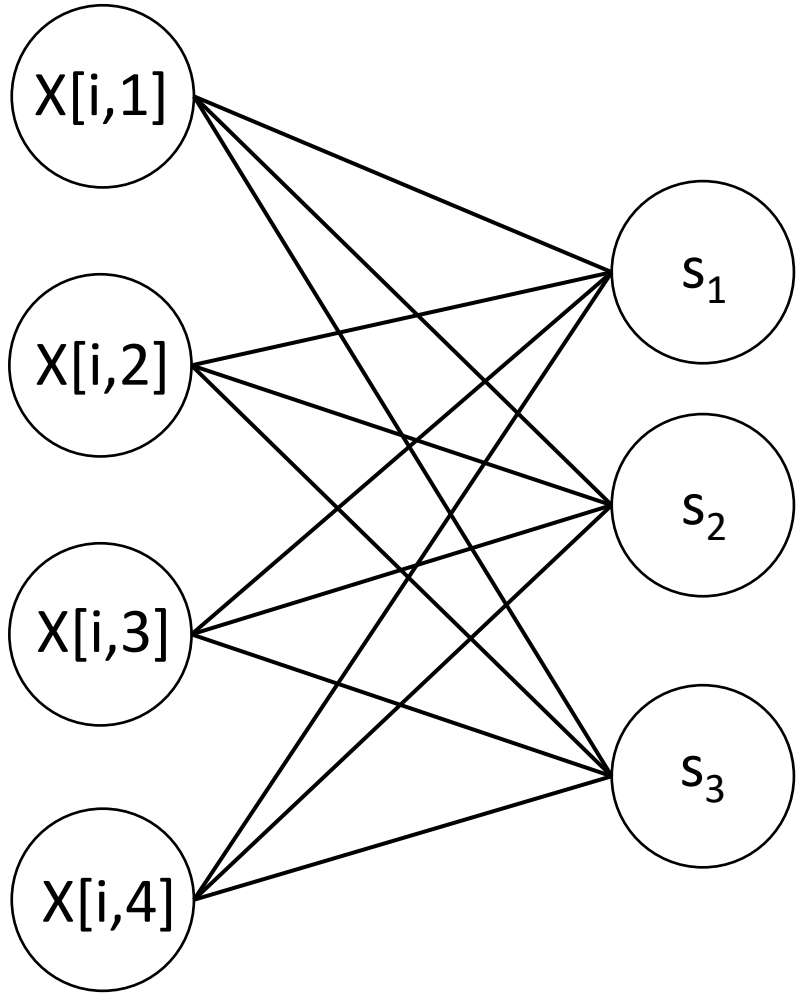
Predicted Labels

If the activation $\boldsymbol{\sigma}$ is the softmax function, then:

$$\hat{y} = \underset{\forall i}{\operatorname{argmax}}(s_i)$$

UNIVERSITY OF CALGARY

Single-layer FCNN



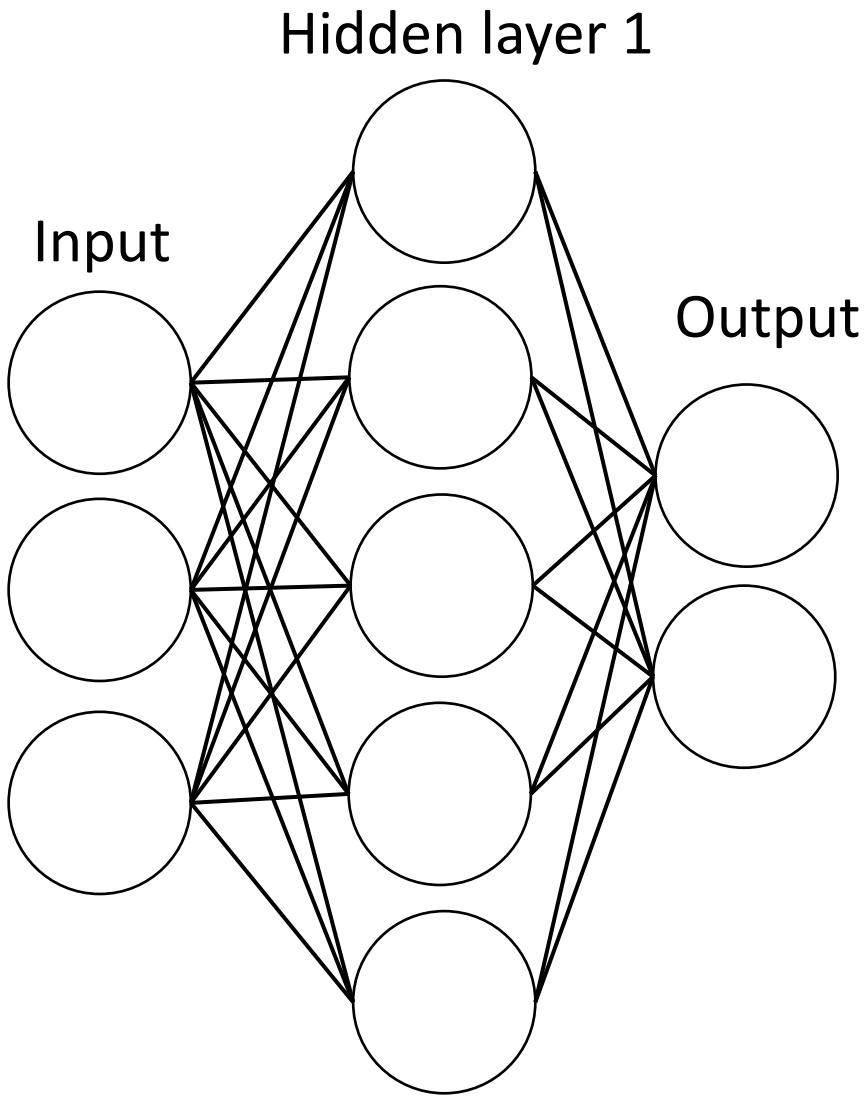
Matrix formulation:

$$[S]_{C\times 1} = \sigma([W]_{C\times M} X^T_{i,:} + [B]_{C\times 1})$$

Number of parameters:

(M + 1) x C = (4 + 1) x 3 = 15

Multi-layer FCNN



$$[S^{(1)}] = \sigma_1([W^{(1)}]X_{i,:}^T + [B^{(1)}])$$

$$[S^{(2)}] = \sigma_2([W^{(2)}]S^{(1)} + [B^{(2)}])$$

Number of parameters:

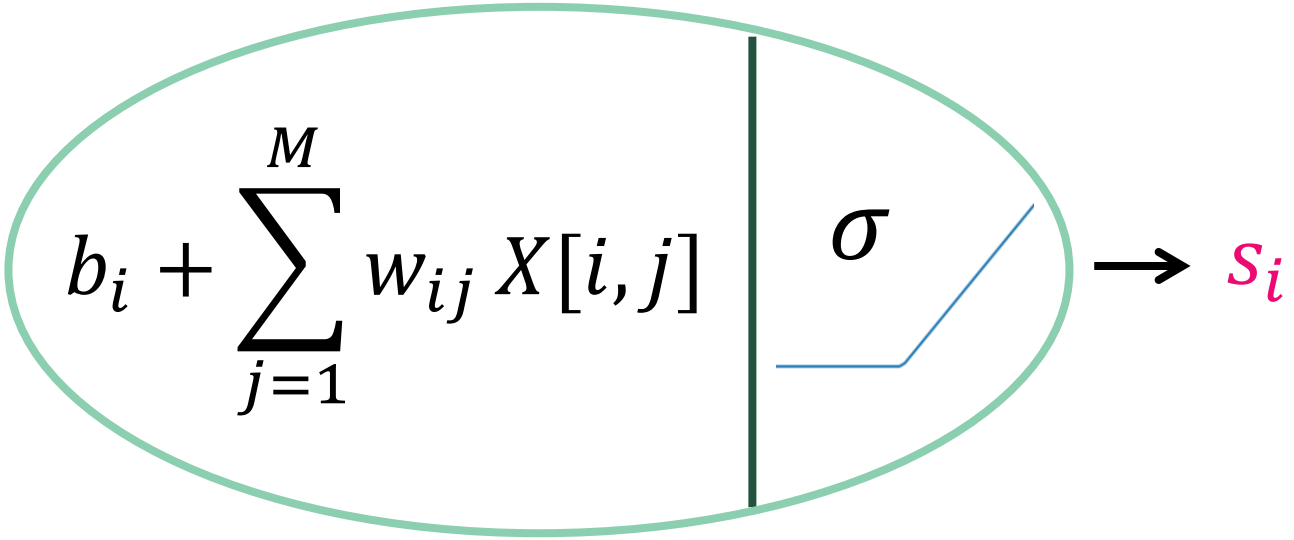
First layer: (3 + 1) x 5 = 20

Second layer: (5 + 1) x 2 = 12

Total: 32

Summary

- Fully connected neural networks alternate linear operations (matrix multiplication + bias term) and non-linear activations



- The number of parameters in each layer is given by the (number of inputs +1) x the number of outputs

Thank you!

UNIVERSITY OF CALGARY