Foundations of Machine Learning (CS 725)

FALL 2024

Lecture 12:

- Introduction to Neural Networks
- Backpropagation

Instructor: Preethi Jyothi

Project Abstracts

Triangulated Attention for Multimodal Learning

Music Genre Prediction

Speech to Speech Translation

Taxi Fare Prediction

Parameter identification and convergence in adaptive control

Sales Forecasting using ML

Song Popularity Prediction

Exploring Vision Transformers

Floor Plan Analysis

Intrusion Response Automation in Cybersecurity

Face Recognition

Neural Style Transfer

ML for Sales Prediction

Spoken Language Identification in Long-form Audio

Learning Deep Representations

Review-based Rating System

OCR

Smart Indoor Positioning

Loan Approval Prediction

Pdf-to-text Conversion

Doodle Recognition x2

Recommendation Systems x2

Project Abstracts

Fraud Detection

Retina Classification

Microplastic Detection in Soil

Modeling of CO2 emissions

Soil Moisture Detection

Predicting Music Hits

Optimising Credit Card Rewards

Pdf to LaTeX

Network Intrusion Detection

Solving Sudoku

Image Segmentation

Fixing Backpropagation by not using it at all

Stock Price Prediction

x3

Multilingual OCR and Translation

Predicting Loan Sanction Decisions

Enhancing Image Resolution

Melody Generation

Multimodal Representation Learning

Organ Segmentation

Radio Isotope Identification

Nifty Price Prediction

Algorithmic Trading

Learning with Noisy Labels

Anime Recommendation

Music Recommendation System

Adaptive Learning Tool

Explainable Recommendation System

2D Image to 3D Model

What is deep learning?

"Deep learning allows computational models that are composed of multiple processing layers to learn representations of data with multiple levels of abstraction."

"Representation learning is a set of methods that allows a machine to be fed with raw data and to automatically discover the representations needed for detection or classification. Deep-learning methods are representation-learning methods with multiple levels of representation, obtained by composing simple but non-linear modules that each transform the representation at one level (starting with the raw input) into a representation at a higher, slightly more abstract level."

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436

History of (Deep) Neural Networks

- McCulloch-Pitts Neuron Model (1943)
- Perceptrons (1957)
- Backpropagation (1960)
- Backpropagation for neural networks (1986)
- Convolutional neural networks (1989)

⋮

- Deep learning for speech recognition (2009)
- AlexNet (2012)
- Generative Adversarial Networks (GANs) (2014)
- AlphaGo (2016)
- Transformers (2018)
- ChatGPT (2022)

Why the resurgence?

- McCulloch-Pitts Neuron Model (1943)
- Perceptrons (1957)
- Backpropagation (1960)
- Backpropagation for neural networks (1986)
- Convolutional neural networks (1989)

⋮

- Deep learning for speech recognition (2009)
- AlexNet (2012)
- Generative Adversarial Networks (GANs) (2014)
- AlphaGo (2016)
- Transformers (2018)
- ChatGPT (2022)

Vast amounts of data

+

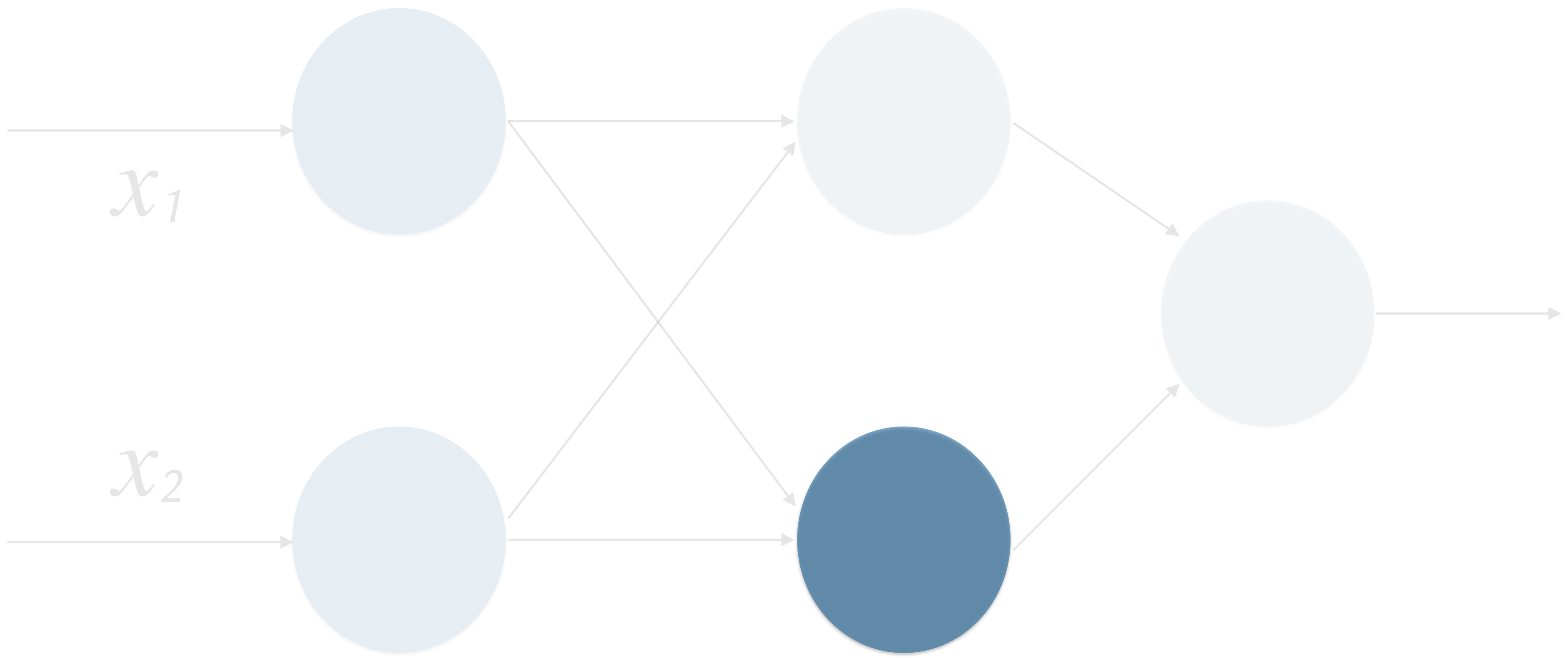
Specialized hardware,
Graphics Processing Units (GPUs)

+

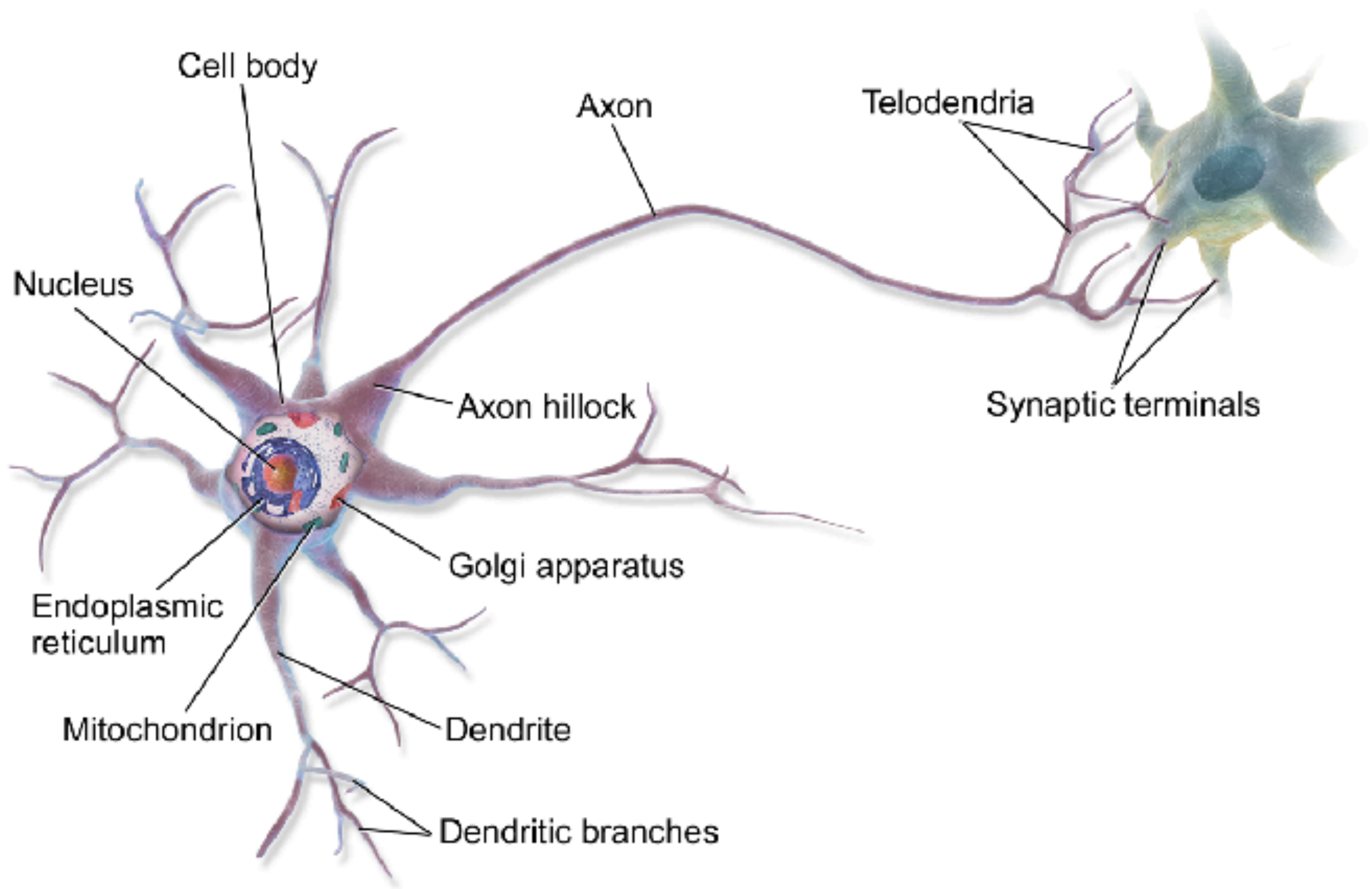
Improved optimization techniques
and new model variants/libraries/toolkits

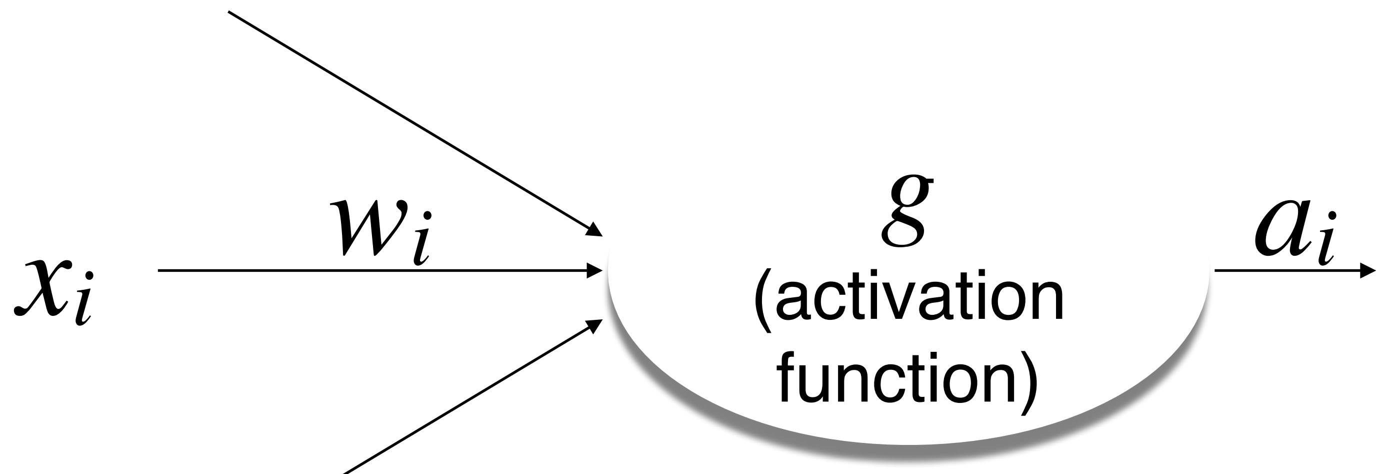
Feed-forward Neural Network

Single Neuron



Single neuron



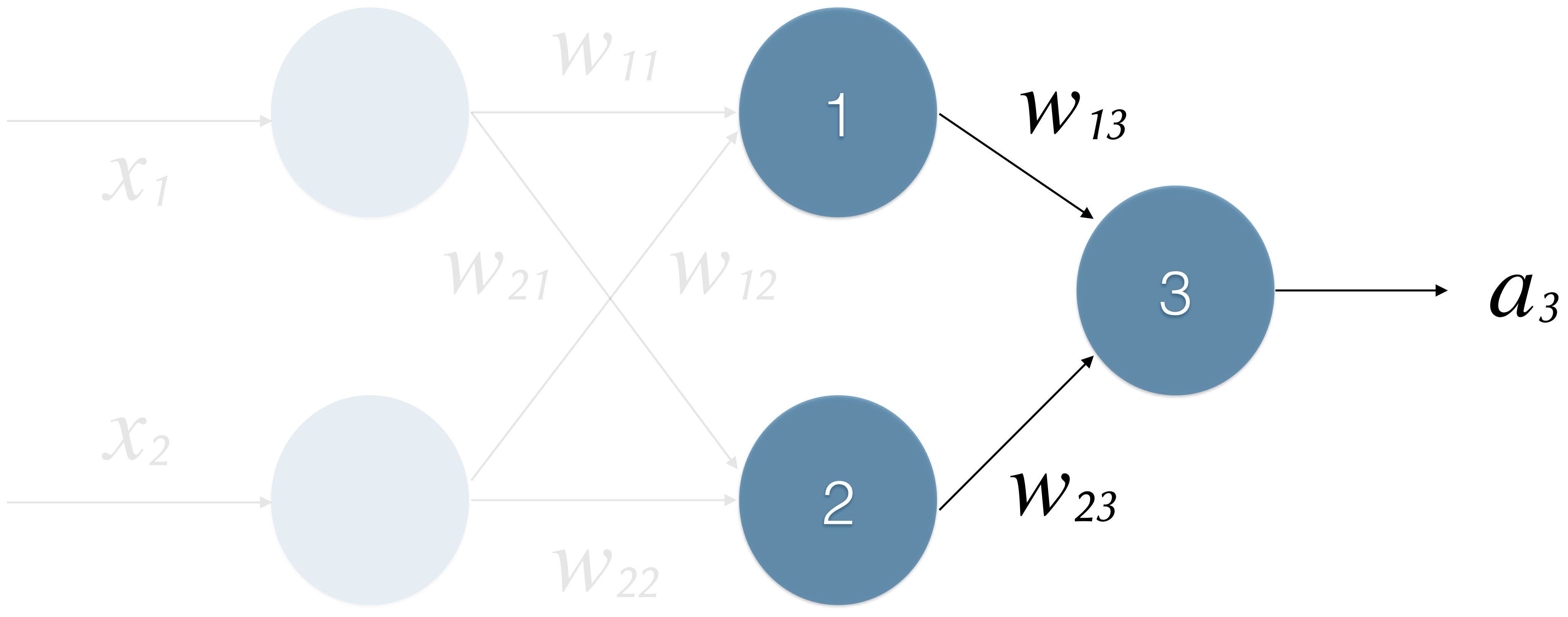


$$a_i = g(\sum_i w_i x_i)$$

Image from: https://upload.wikimedia.org/wikipedia/commons/1/10/Blausen_0657_MultipolarNeuron.png

Feed-forward Neural Network

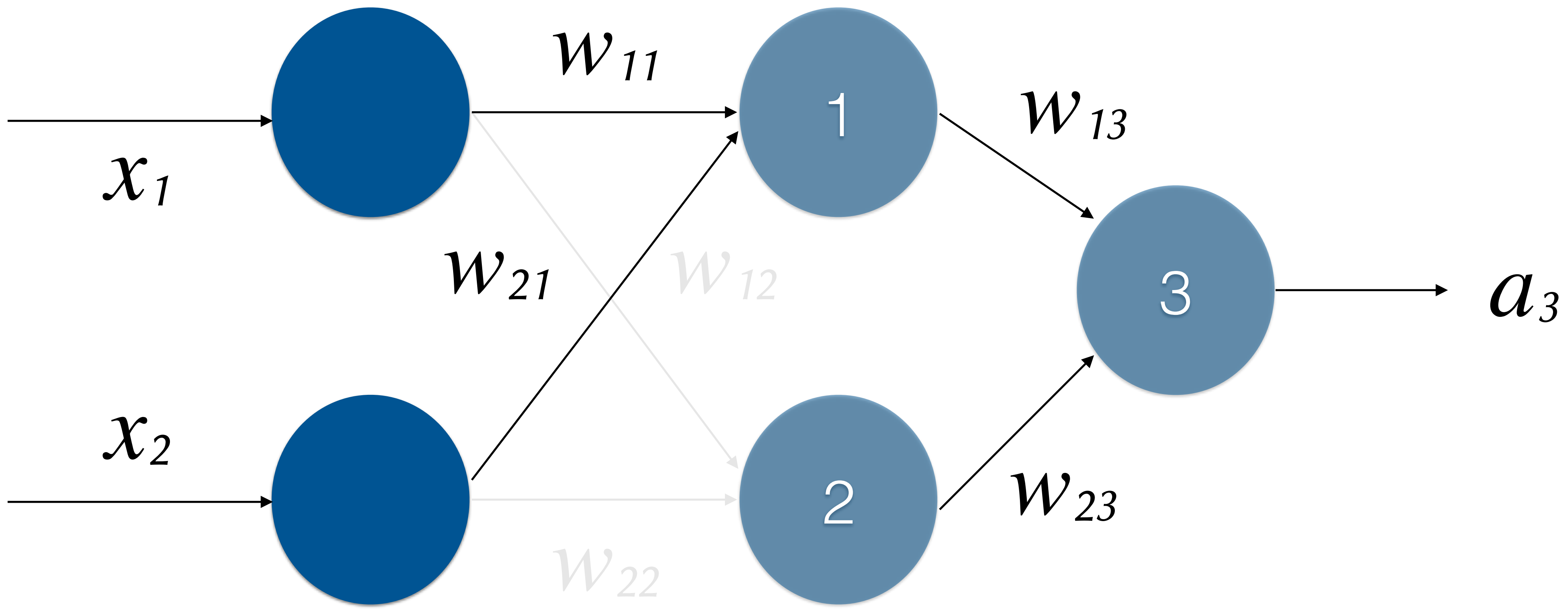
Parameterized Model



$$a_3 = g(w_{13} \cdot a_1 + w_{23} \cdot a_2 + b_3)$$

Feed-forward Neural Network

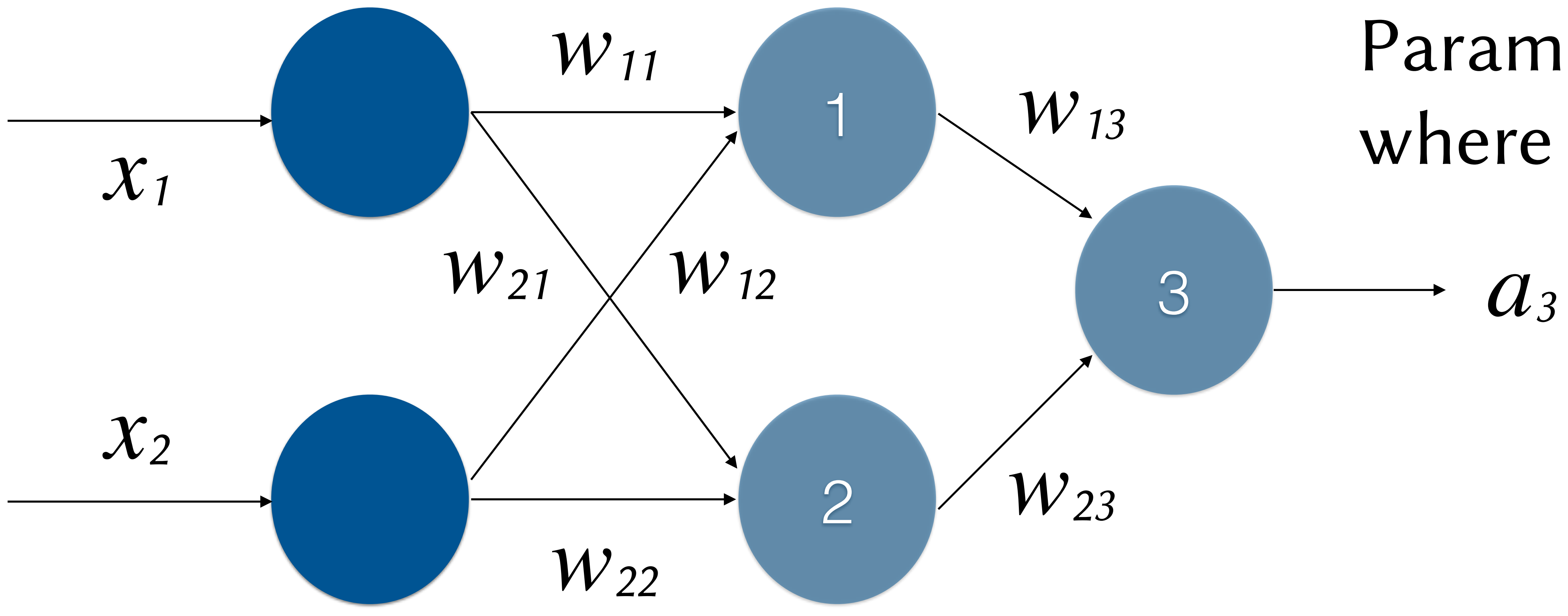
Parameterized Model



$$
\begin{aligned}
a_3 &= g(w_{13} \cdot a_1 + w_{23} \cdot a_2 + b_3) \\
&= g(w_{13} \cdot \big(g(w_{11} \cdot x_1 + w_{21} \cdot x_2 + b_1)\big) \\
&\qquad\qquad + \qquad \cdots
\end{aligned}
$$

Feed-forward Neural Network

Parameterized Model



Parameters of the network: $\mathrm{NN}(\mathbf{x};\theta)$ where θ refers to all w_{ij} and b_j

$$
\begin{aligned}
a_3 &= g(w_{13} \cdot a_1 + w_{23} \cdot a_2 + b_3) \\
&= g(w_{13} \cdot \big(g(w_{11} \cdot x_1 + w_{21} \cdot x_2 + b_1)\big) \\
&\qquad + \; w_{23} \cdot \big(g(w_{12} \cdot x_1 + w_{22} \cdot x_2 + b_2)\big) + b_3)
\end{aligned}
$$

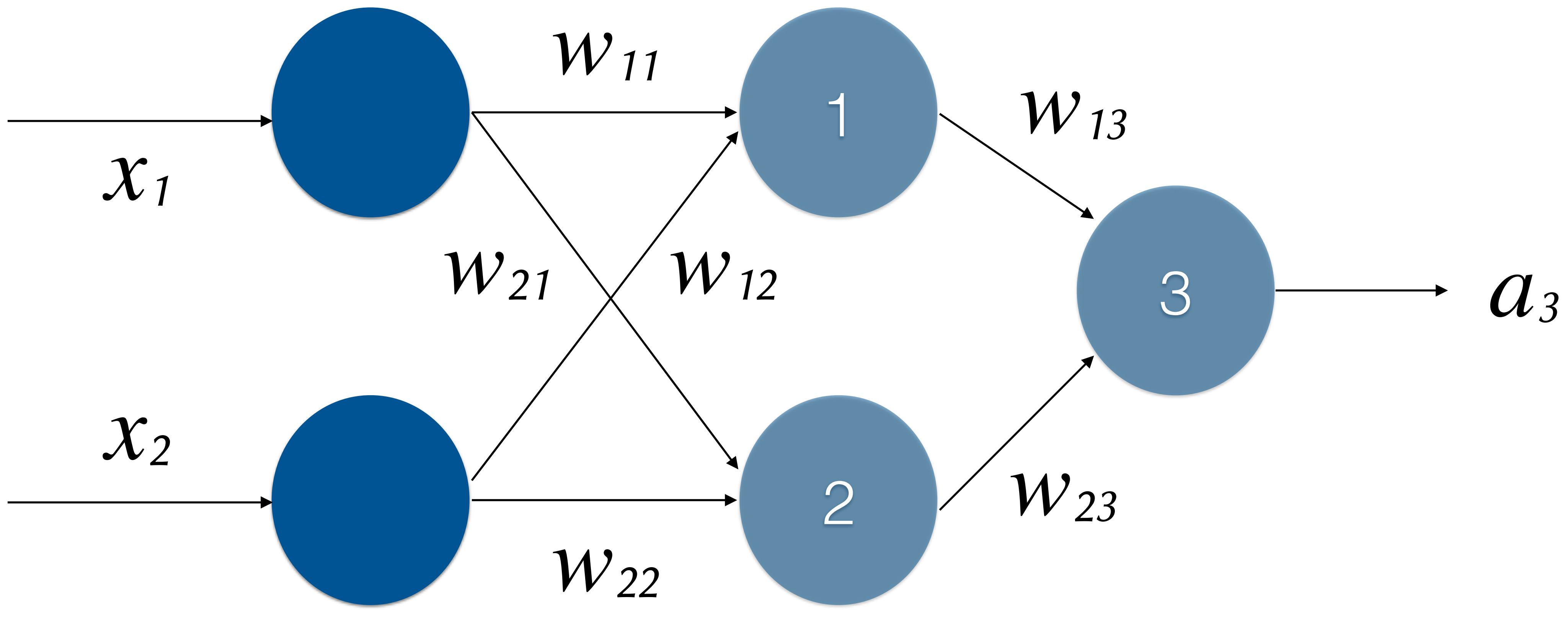
Compact matrix notation: Input $\mathbf{x} = [x_1, x_2]$ is written as a 2-dimensional vector and the layer above it is a 2-dimensional vector $\mathbf{h}$, a fully-connected layer is associated with:

$$\mathbf{h} = \mathbf{x}\mathbf{W} + \mathbf{b}$$

where w_{ij} in $\mathbf{W}$ is the weight of the connection between j^{th} neuron in the input row and i^{th} neuron in the first hidden layer and $\mathbf{b}$ is the bias vector.

Feed-forward Neural Network

Parameterized Model



$$a_3 = g(w_{13} \cdot a_1 + w_{23} \cdot a_2 + b_3)$$
$$= g(w_{13} \cdot \big(g(w_{11} \cdot x_1 + w_{21} \cdot x_2 + b_1)\big)$$
$$+ \ w_{23} \cdot \big(g(w_{12} \cdot x_1 + w_{22} \cdot x_2 + b_2)\big) + b_3)$$

The simplest neural network is the perceptron:

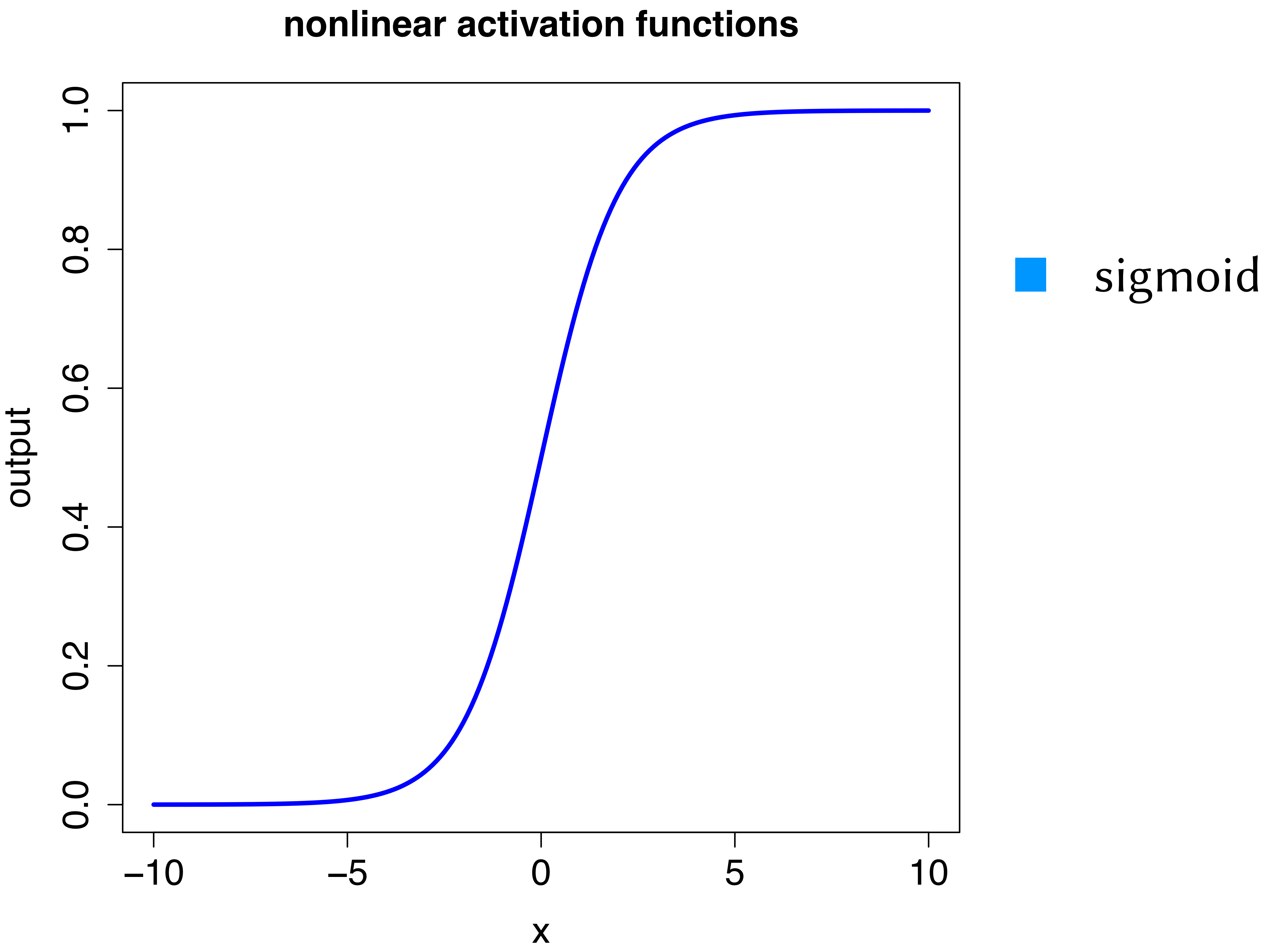
$$\text{Perceptron}(\mathbf{x}) = \mathbf{x}\mathbf{W} + \mathbf{b}$$

A 1-layer feedforward neural network (multi-layer perceptron) has the form:

$$\text{MLP}(\mathbf{x}) = g(\mathbf{x}\mathbf{W}_1 + \mathbf{b}_1)\mathbf{W}_2 + \mathbf{b}_2$$

Common Activation Functions (g)

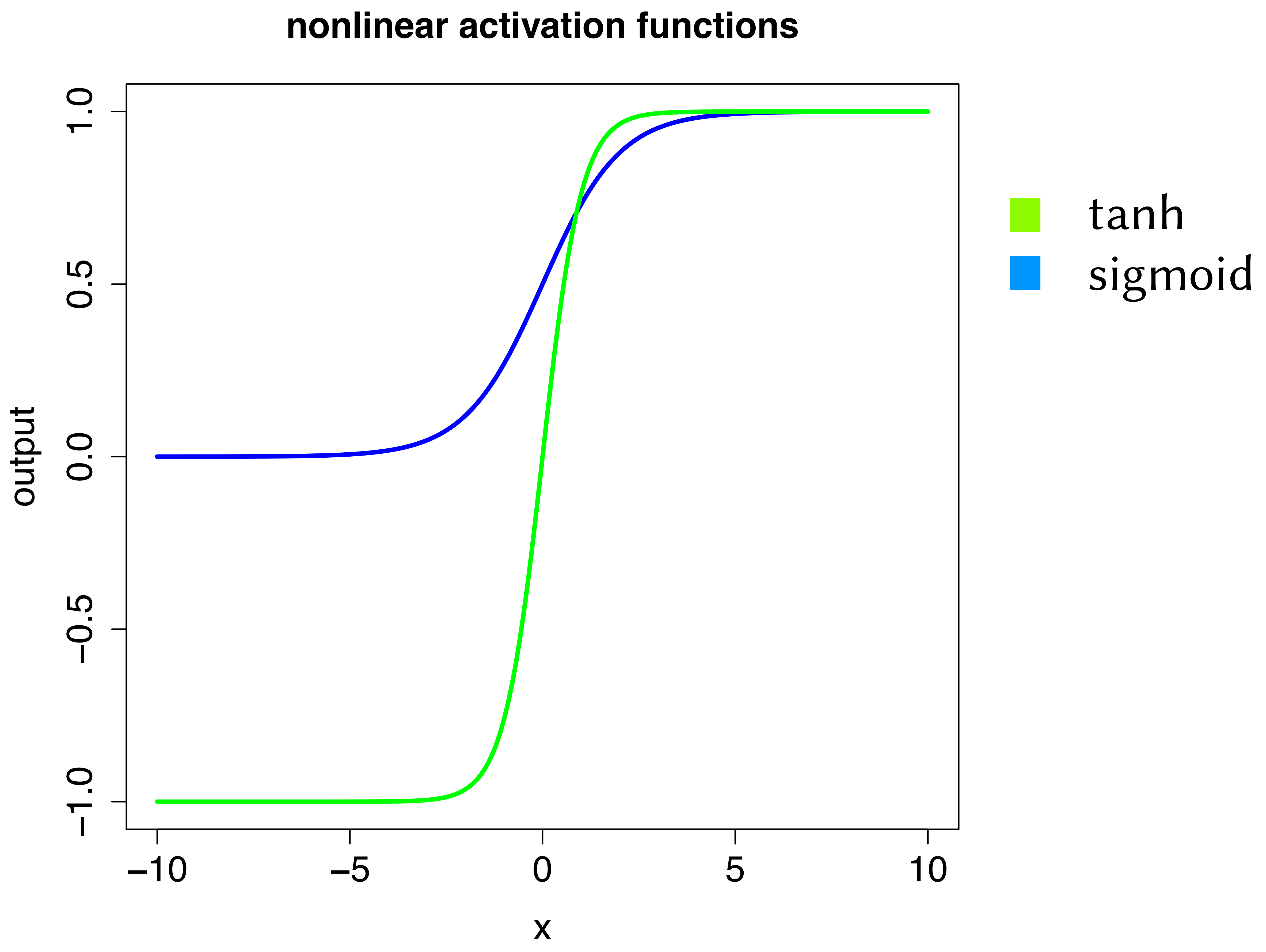
Sigmoid: $\sigma(x) = 1/(1 + e^{-x})$



Common Activation Functions (g)

Sigmoid: $\sigma(x) = 1/(1 + e^{-x})$

Hyperbolic tangent (tanh): $\tanh(x) = (e^{2x} - 1)/(e^{2x} + 1)$



Common Activation Functions (g)

Sigmoid: $\sigma(x) = 1/(1 + e^{-x})$

Hyperbolic tangent (tanh): $\tanh(x) = (e^{2x} - 1)/(e^{2x} + 1)$

Rectified Linear Unit (ReLU): $\mathrm{RELU}(x) = \max(0, x)$