Introduction to Neural Networks

Agenda

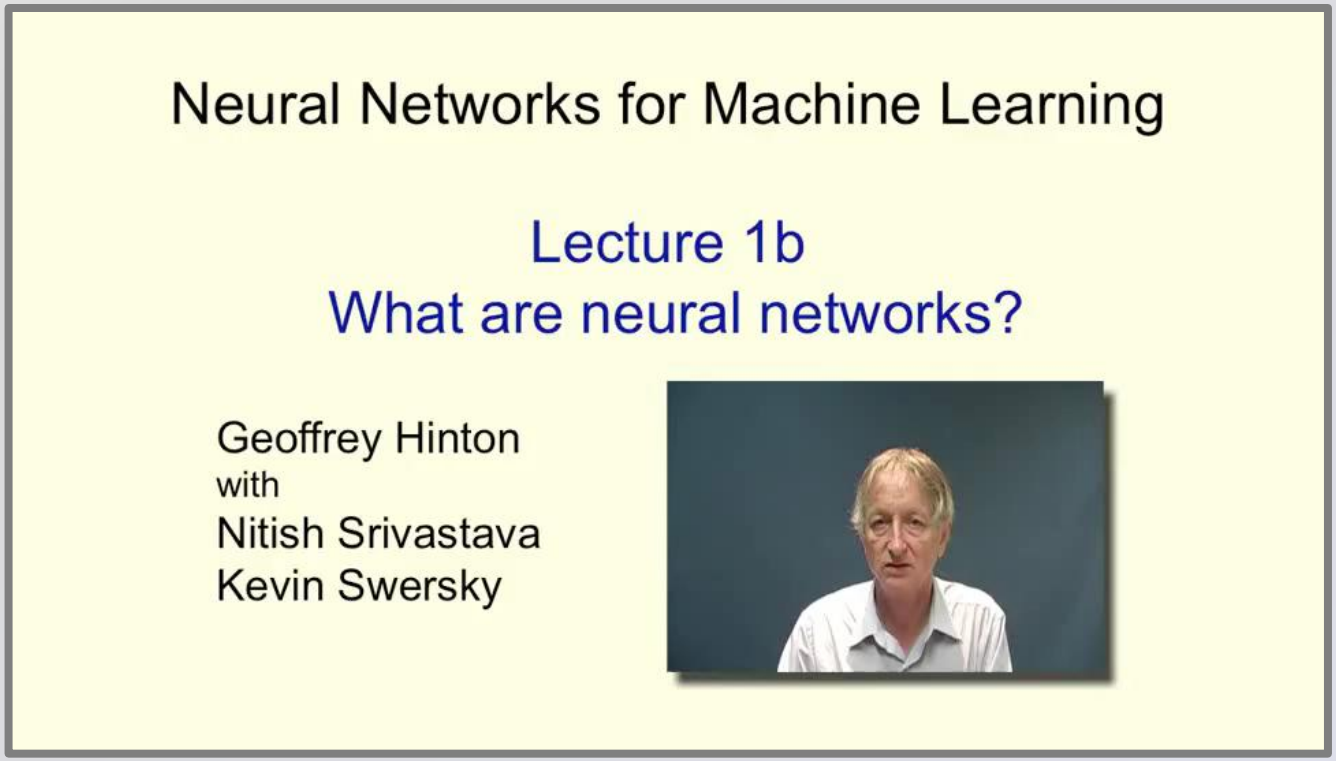
- Introduction to the Brain
- Introduction to Neural Networks
- Gradient Descent
- Resources for Further Study

Agenda

- Introduction to the Brain
- Introduction to Neural Networks
- Gradient Descent
- Resources for Further Study

Inspiration

Geoffrey Hinton, a pioneer in the resurgence of interest in Neural Networks shares an overview of the brain.



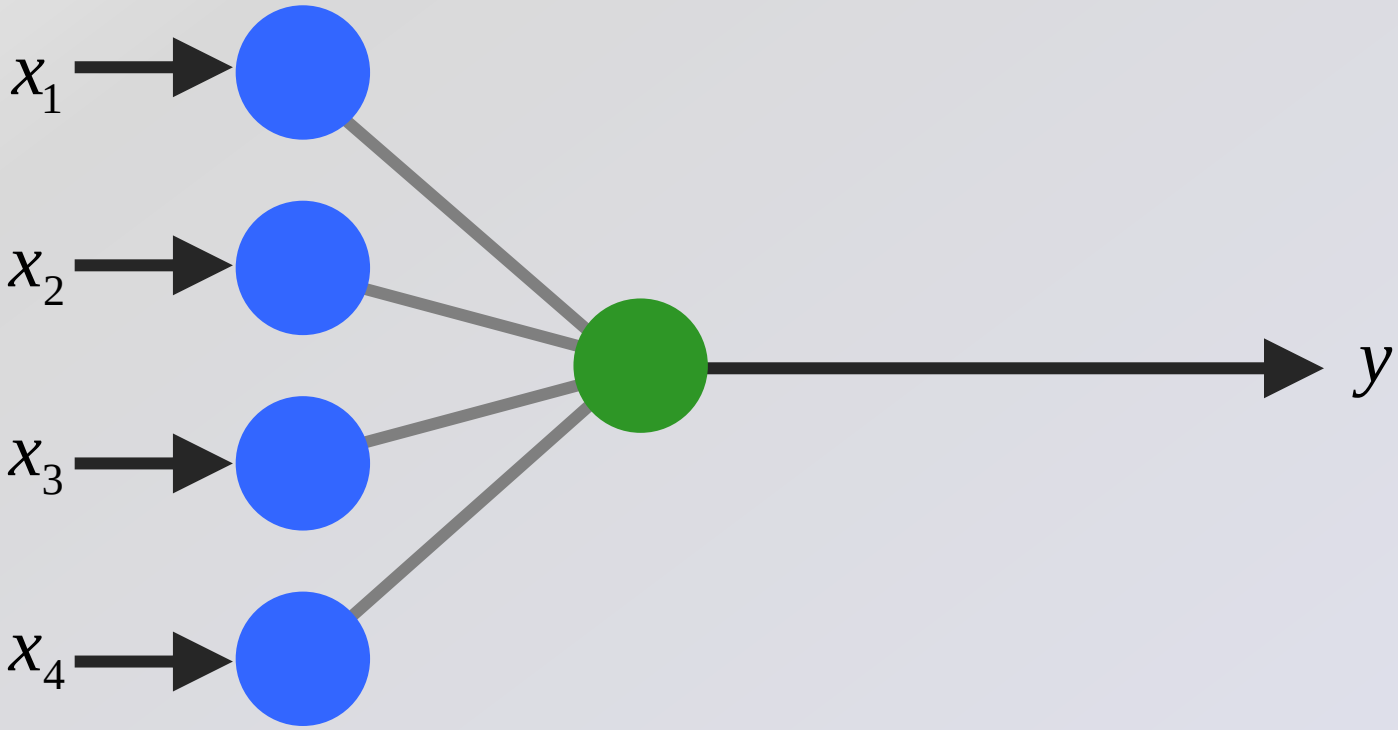
Source: https://class.coursera.org/neuralnets-2012-001/lecture/7

Agenda

- Introduction to the Brain
- Introduction to Neural Networks
- Gradient Descent
- Resources for Further Study

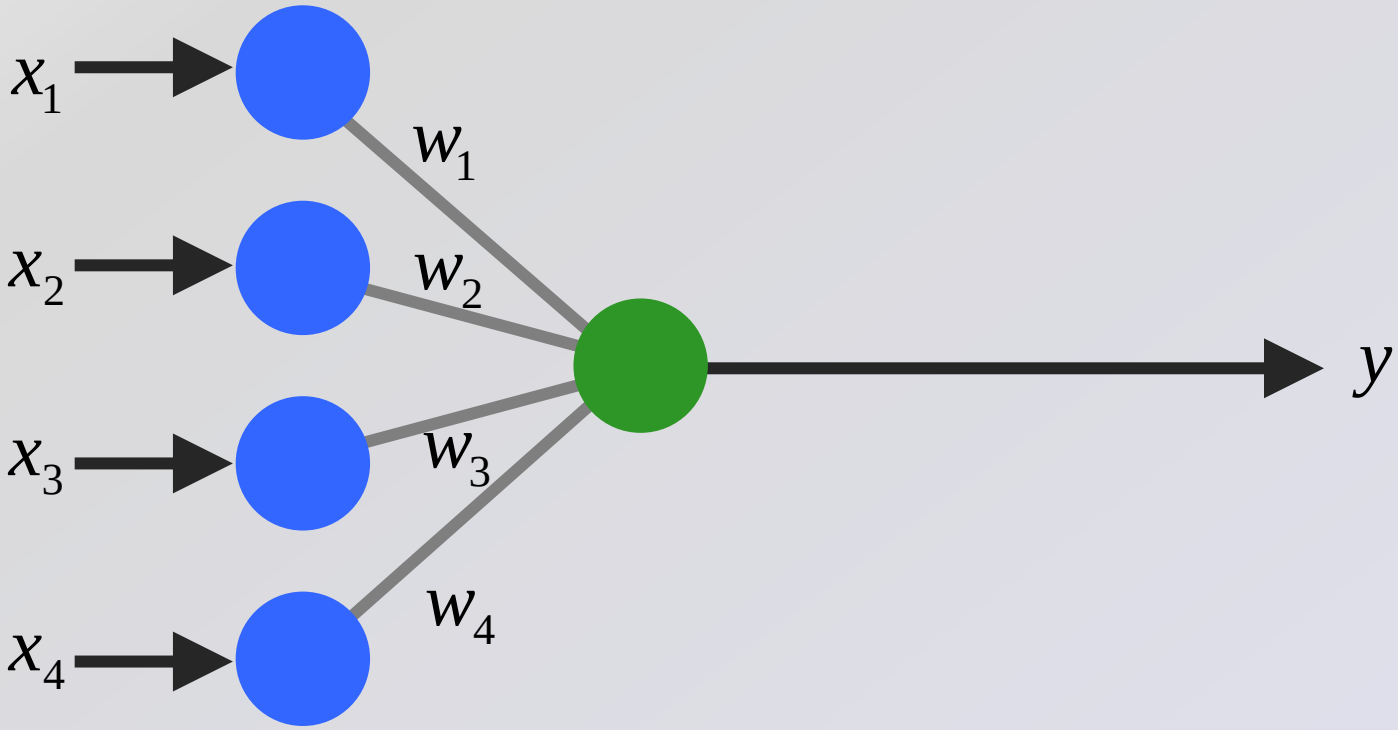
Perceptron

Initially, in the 1960s, artificial neurons were modeled using something called a perceptron, which is a step function



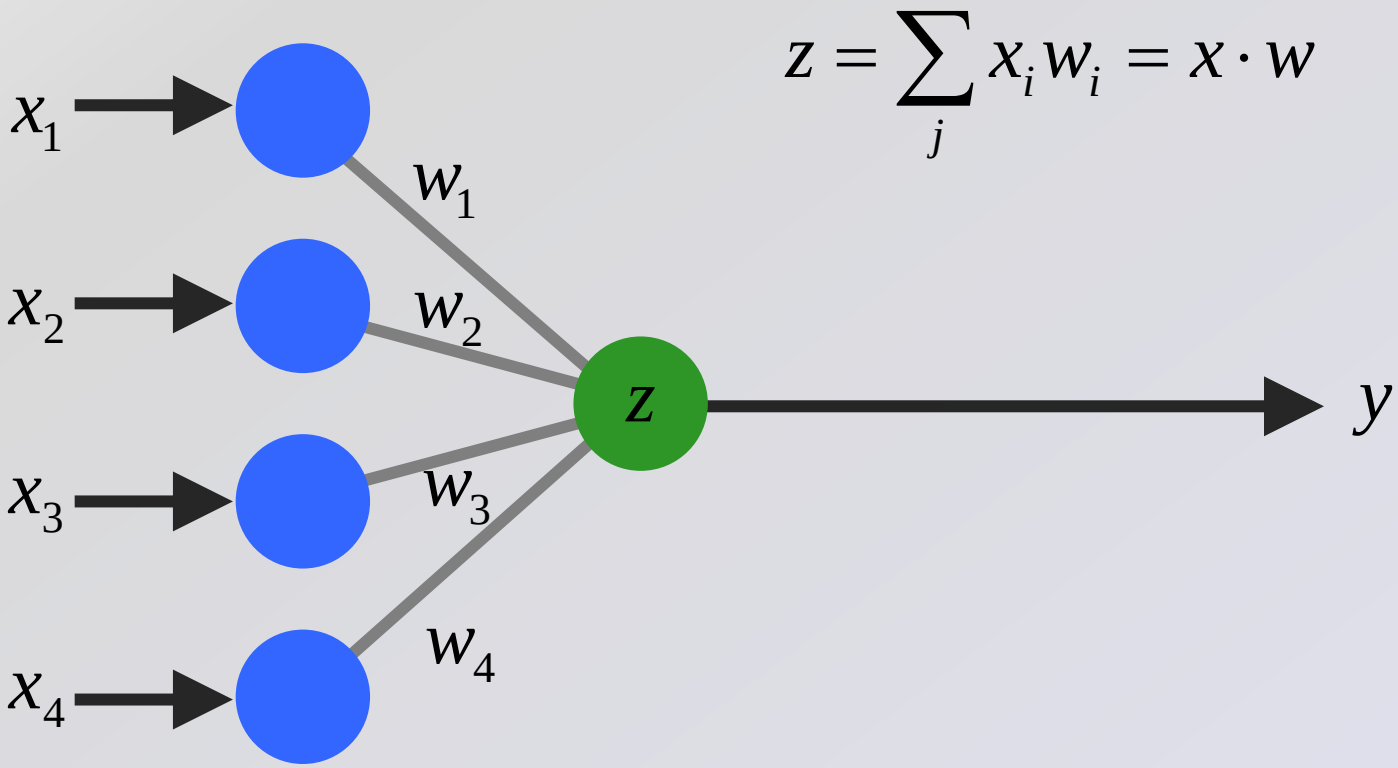
Perceptron

Initially, in the 1960s, artificial neurons were modeled using something called a perceptron, which is a step function



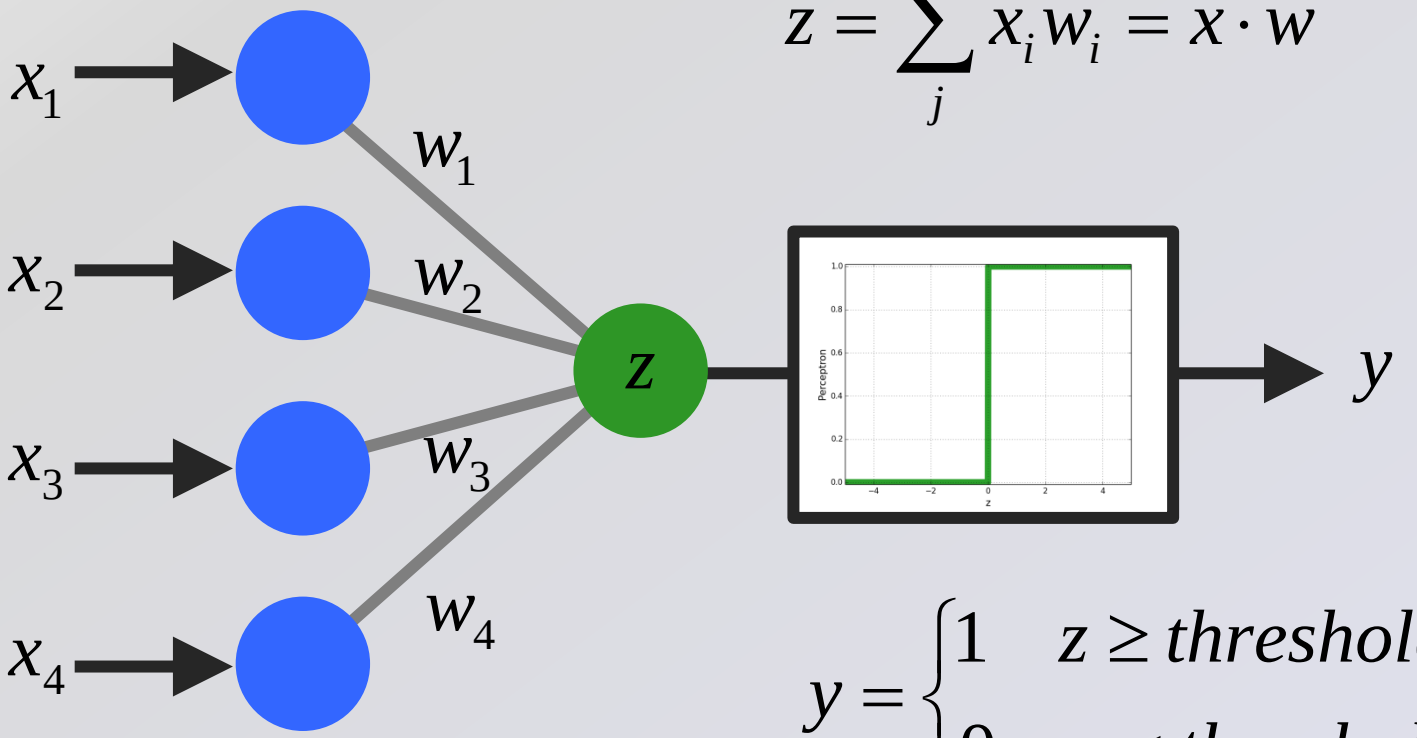
Perceptron

Initially, in the 1960s, artificial neurons were modeled using something called a perceptron, which is a step function



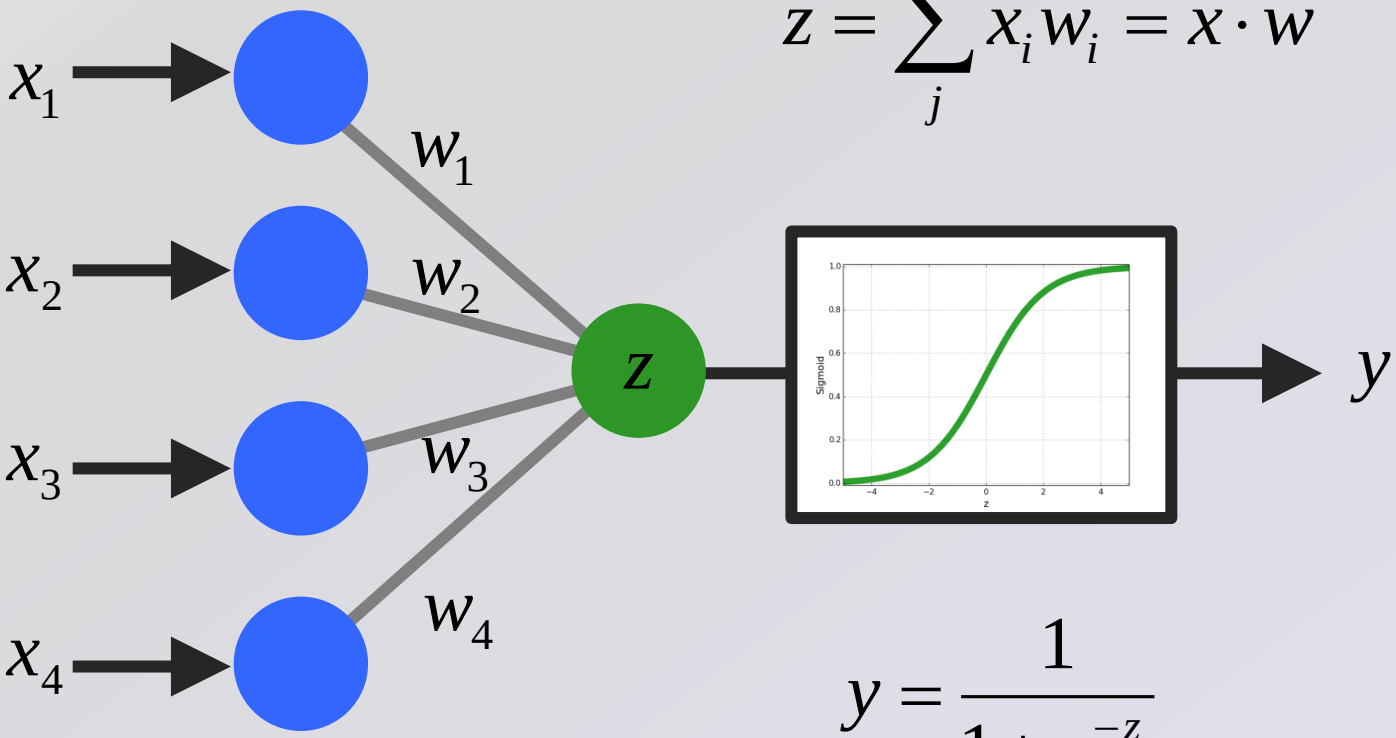
Perceptron

Initially, in the 1960s, artificial neurons were modeled using something called a perceptron, which is a step function



Sigmoid Function

In the 1980s, researchers realized they could get better results from a smooth activation function, such as the sigmoid function.

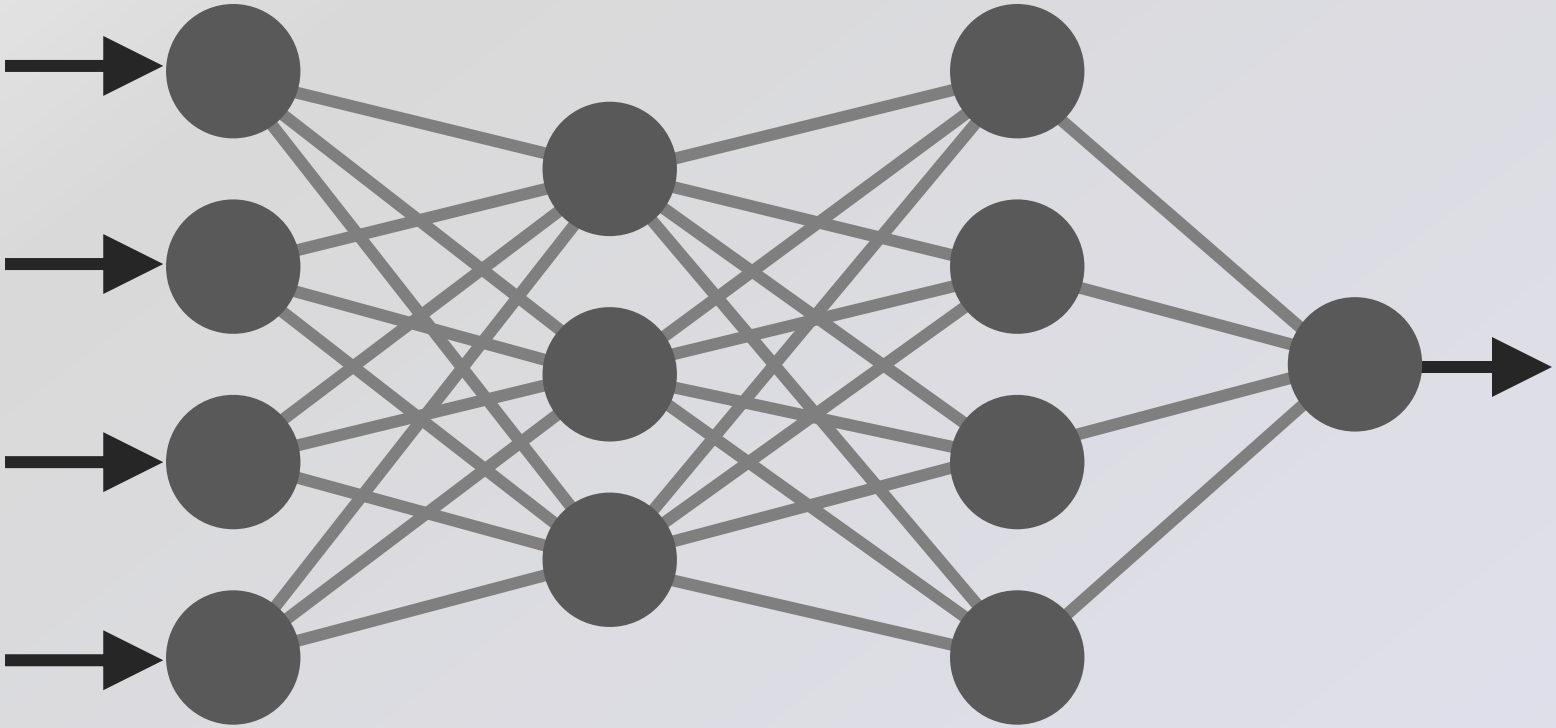


$$z = \sum_j x_i w_i = x \cdot w$$

$$y = \frac{1}{1 + e^{-z}}$$

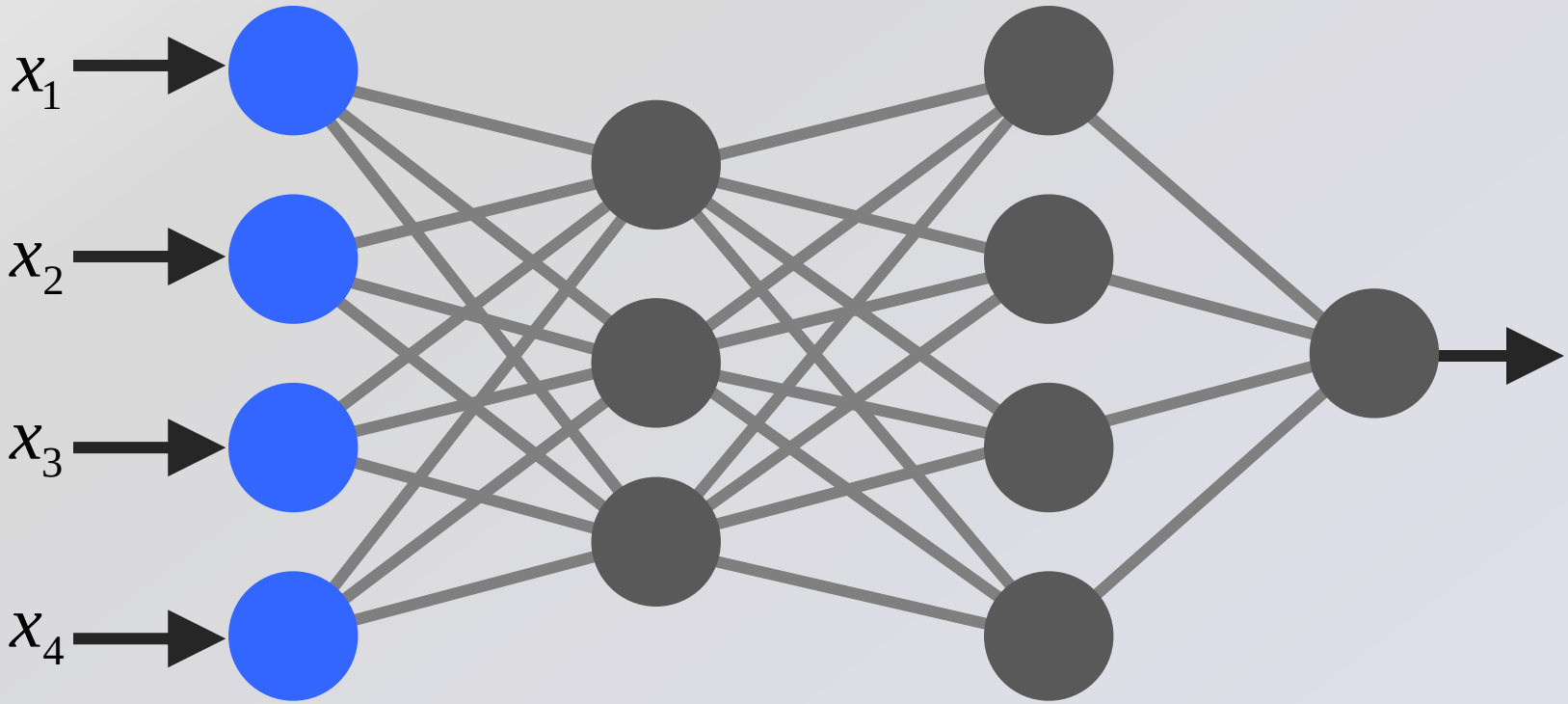
Representation

Artificial neural networks are represented as a system of interconnected neurons with multiple layers.



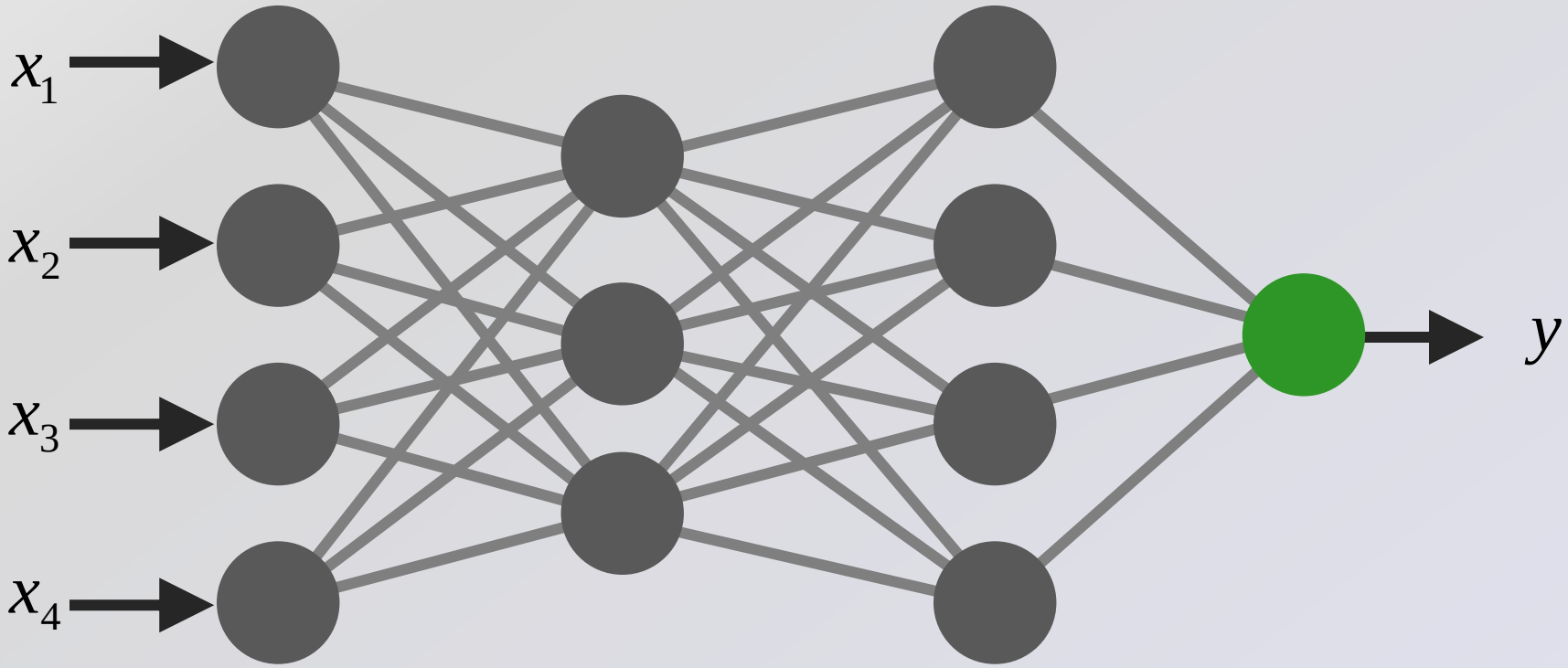
Input Layer

Input neurons represent the features of the dataset.



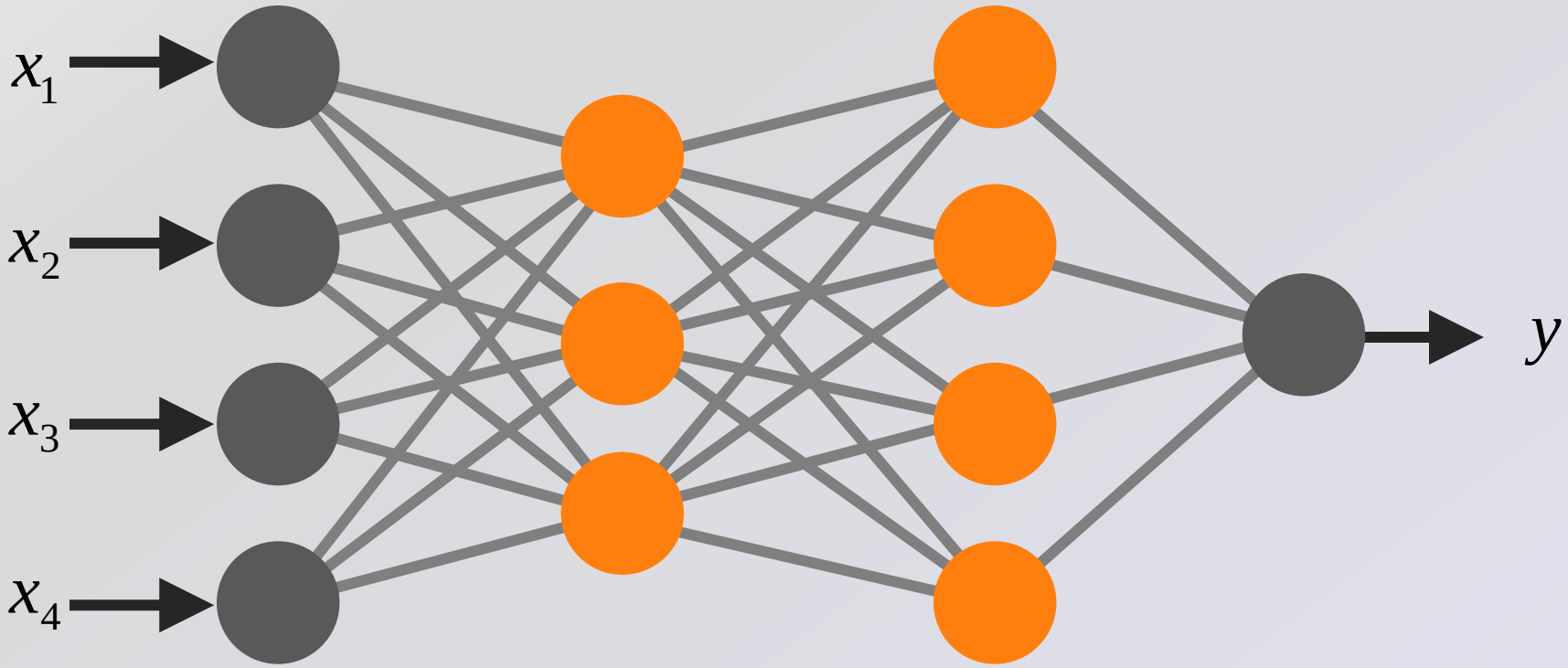
Output Layer

For supervised learning, the output neuron represents the response variable(s).



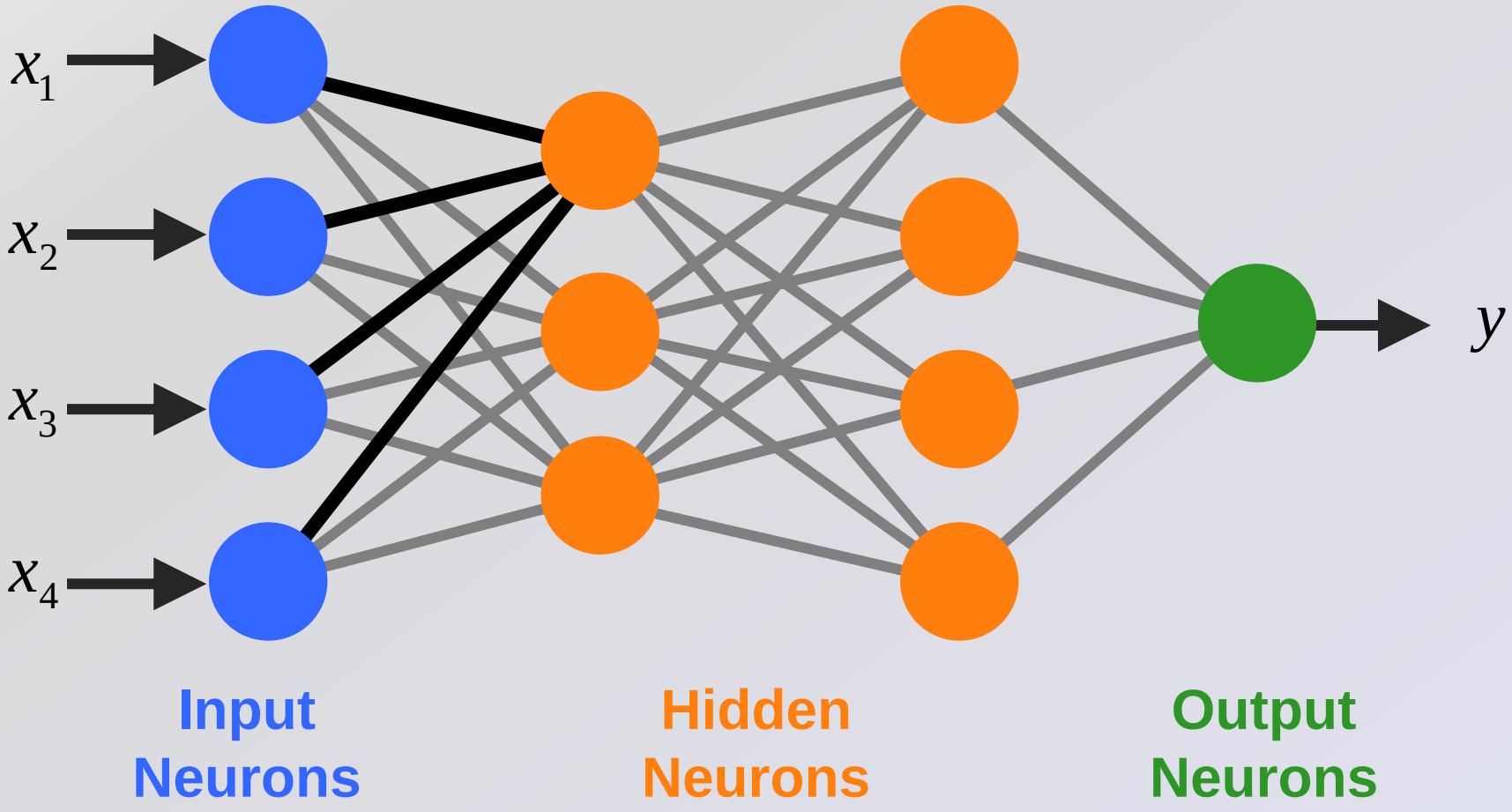
Hidden Layers

Hidden layer are what allow the network to fit highly non-linear patterns.



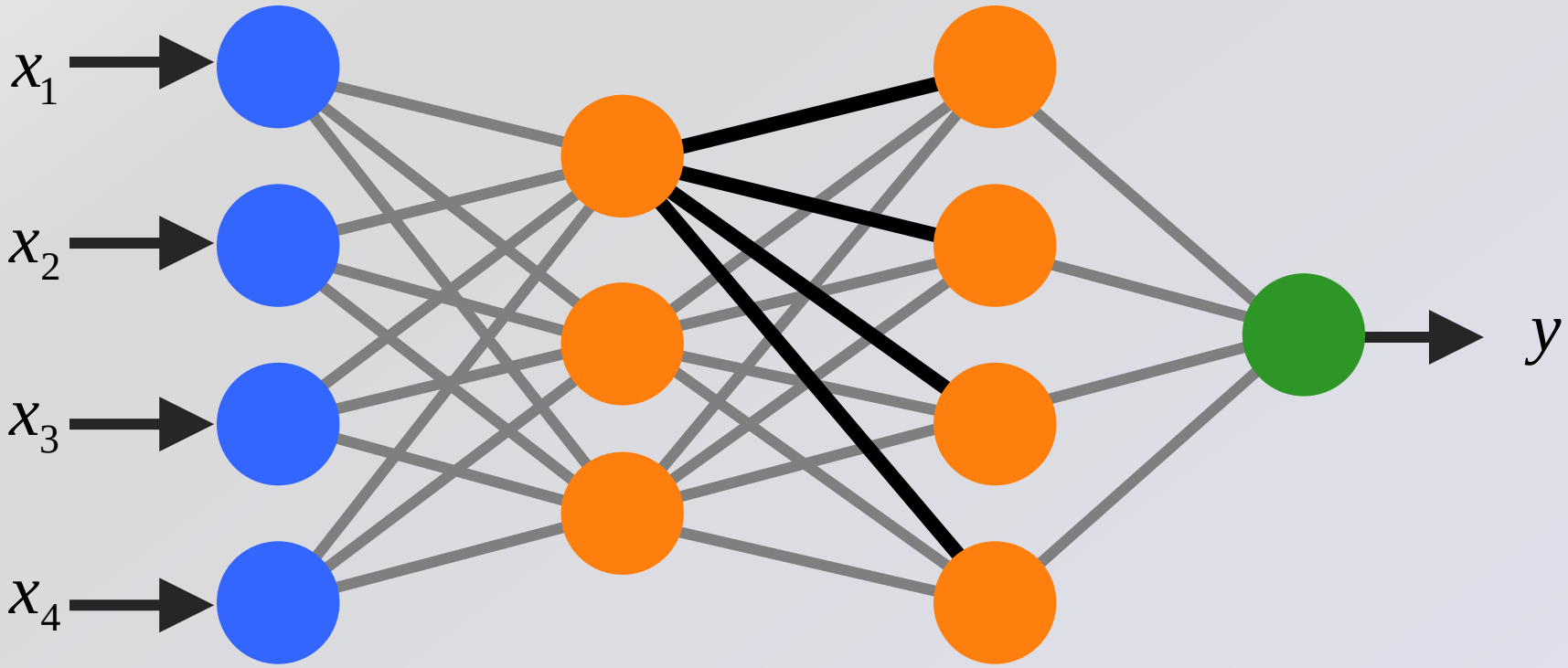
Representation

For feedforward neural networks, the value for each neuron depends only on the neuron values in the previous layer.



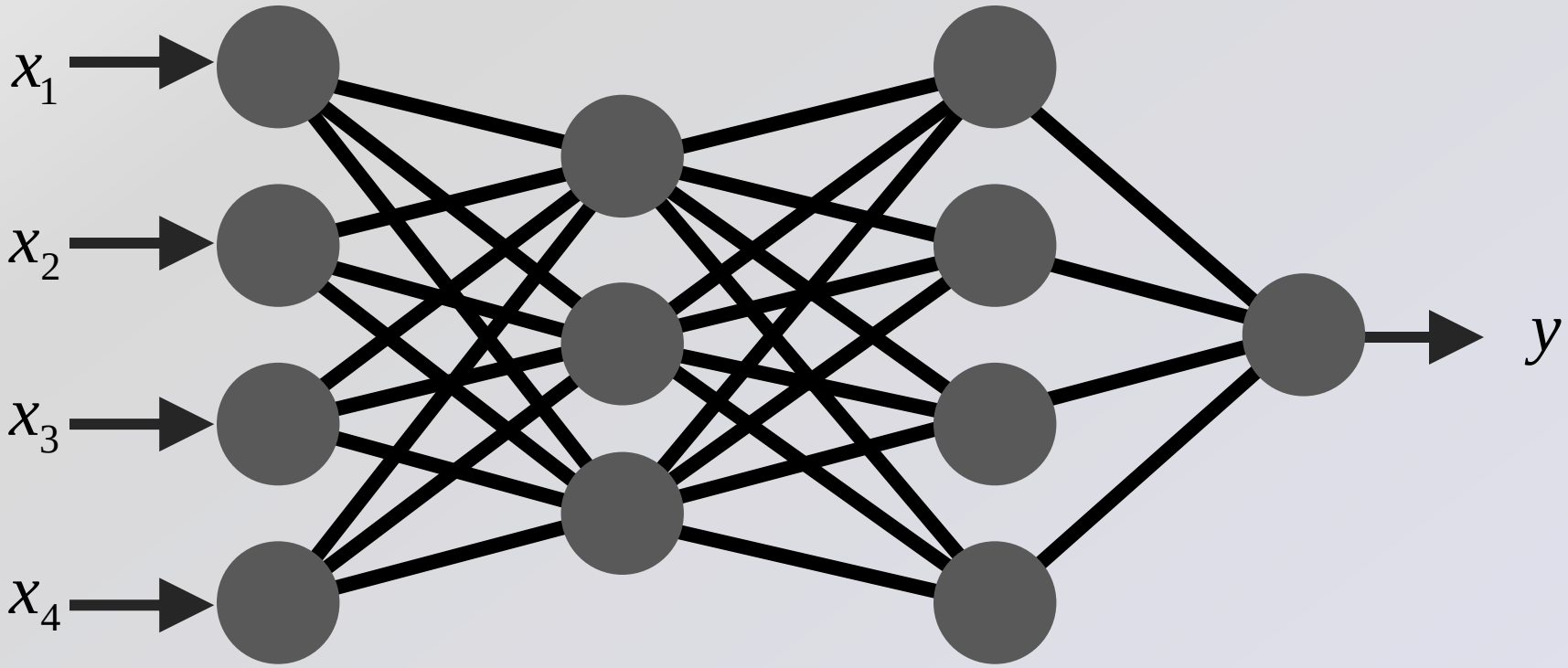
Representation

The value for each neuron influences the neuron values in the subsequent layer.



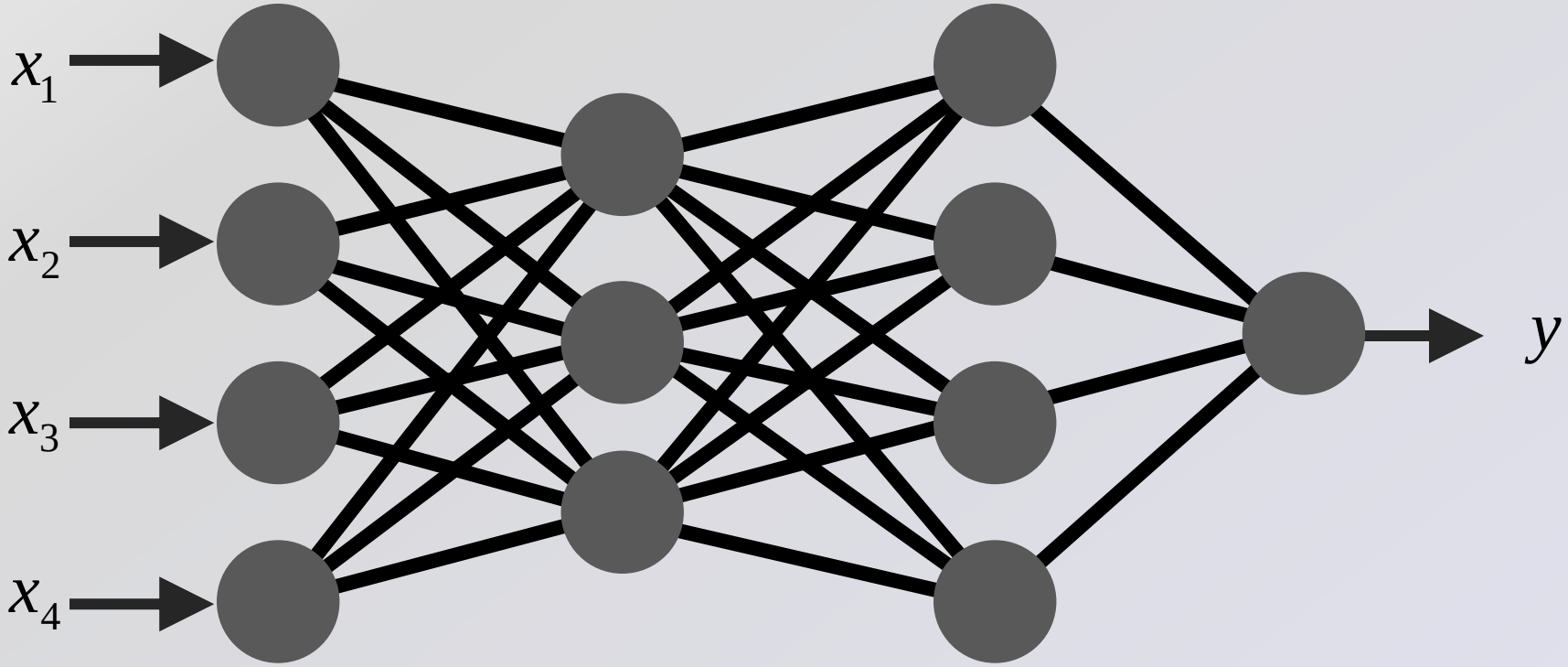
Weights

Connections between nodes are defined by a weight value which determines the influence of one neuron on another.



Weights

These weights are learned through an iterative procedure whereby the change in output is repeatedly calculated for gradual changes in the weights.



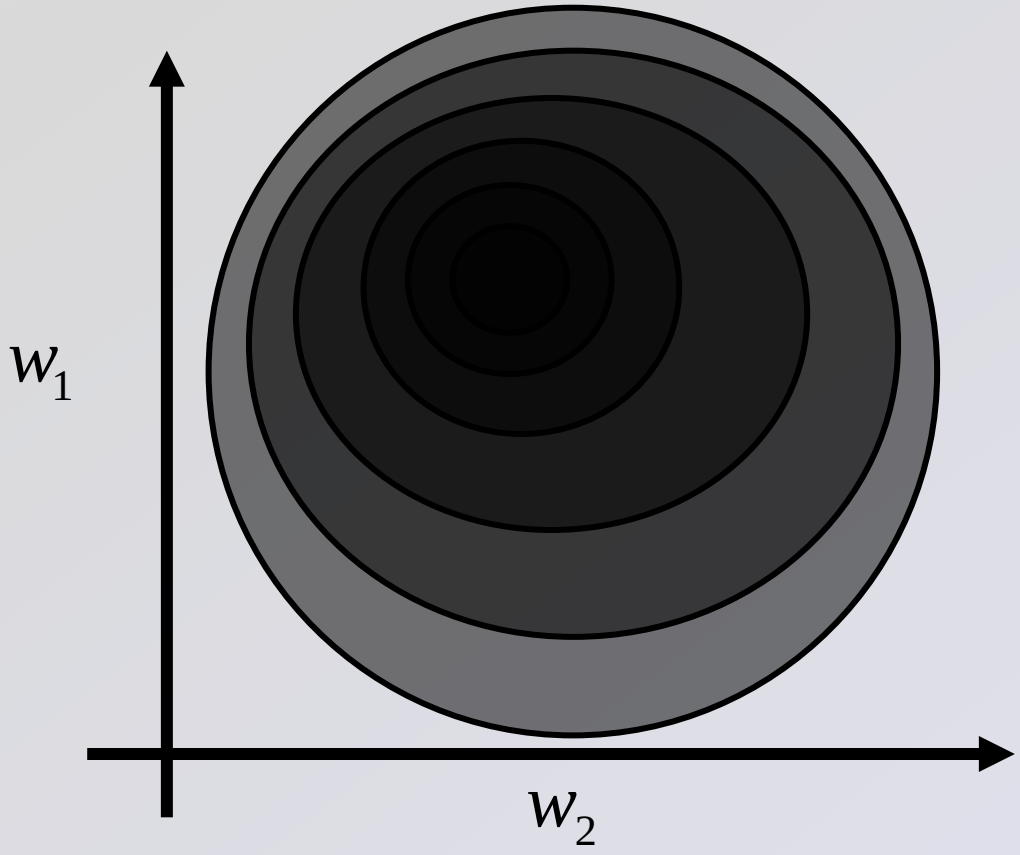
$$w + \Delta w \longrightarrow y + \Delta y$$

Agenda

- Introduction to the Brain
- Introduction to Neural Networks
- Gradient Descent
- Resources for Further Study

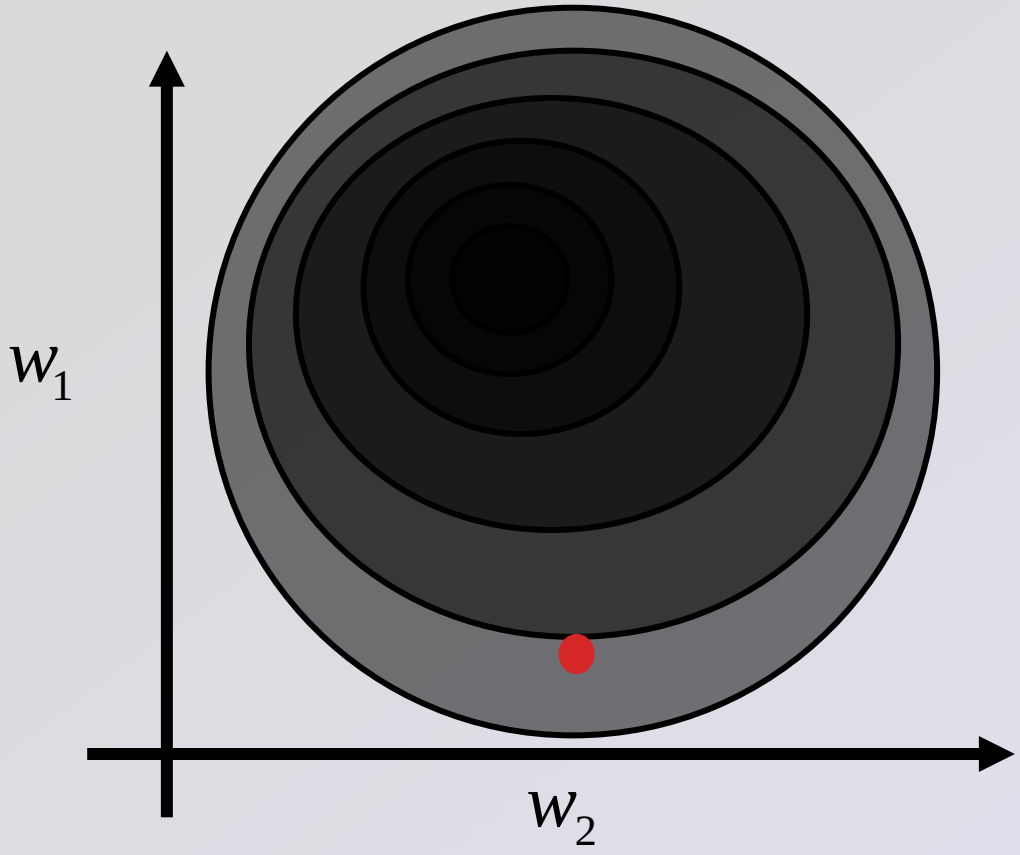
Gradient Descent - Intuition

Weights are updating according to an algorithm called the gradient descent, an iterative procedure for finding the optimal parameters for a given model.



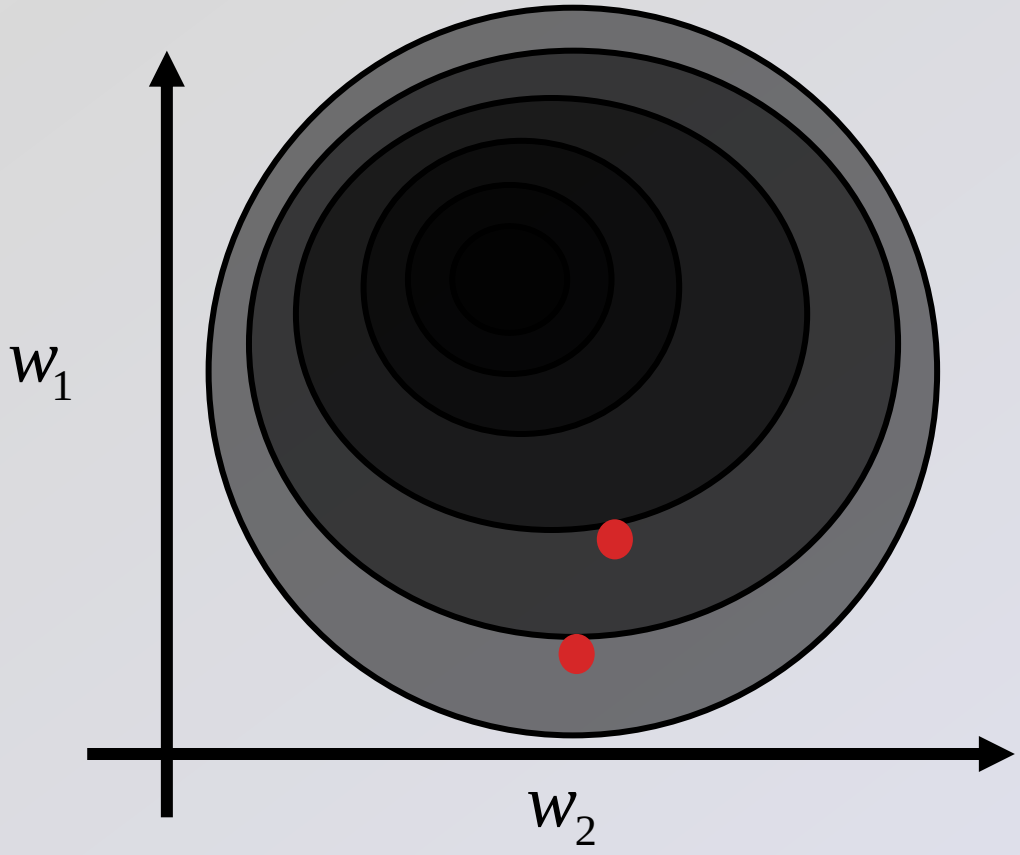
Gradient Descent - Intuition

Weights are updating according to an algorithm called the gradient descent, an iterative procedure for finding the optimal parameters for a given model.



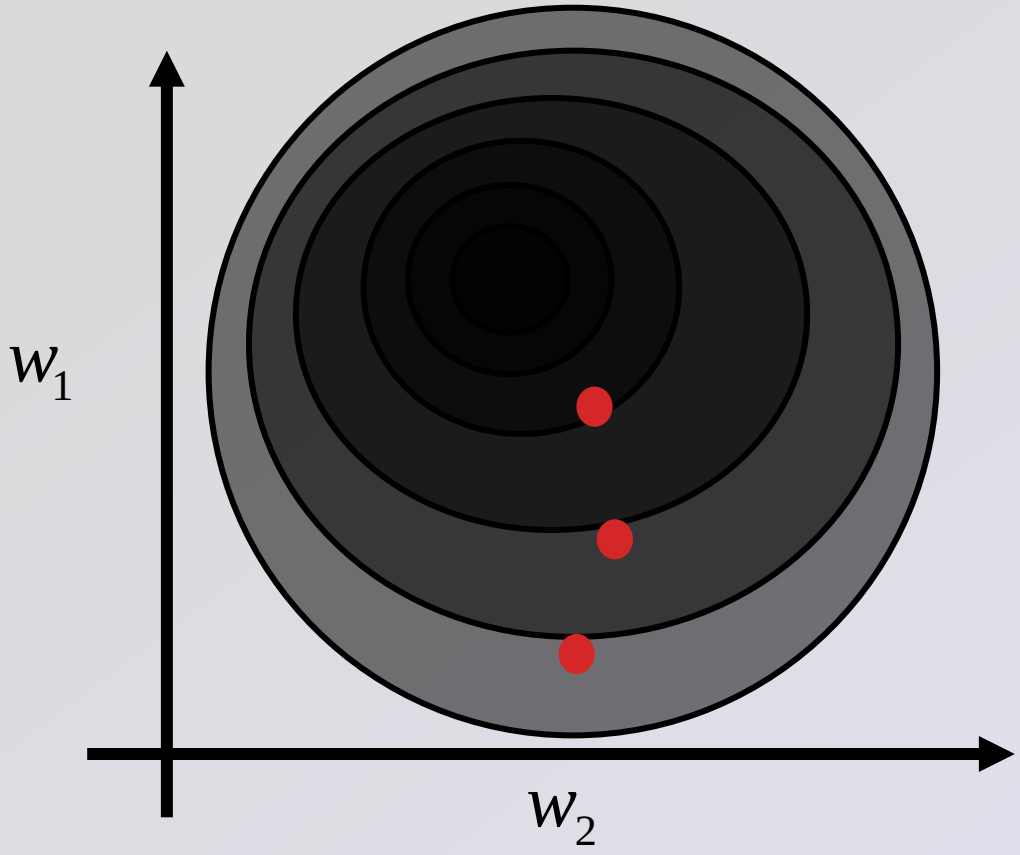
Gradient Descent - Intuition

Weights are updating according to an algorithm called the gradient descent, an iterative procedure for finding the optimal parameters for a given model.



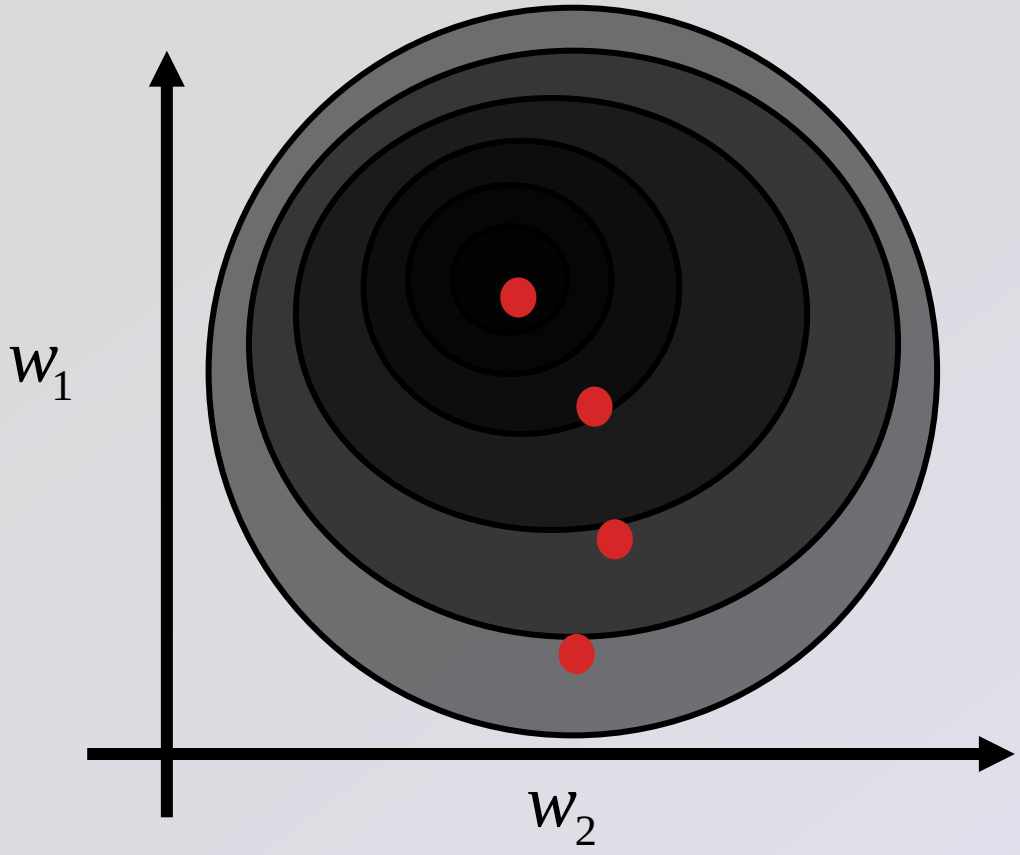
Gradient Descent - Intuition

Weights are updating according to an algorithm called the gradient descent, an iterative procedure for finding the optimal parameters for a given model.



Gradient Descent - Intuition

Weights are updating according to an algorithm called the gradient descent, an iterative procedure for finding the optimal parameters for a given model.



Meta-Evaluation

Meta-Evaluation

Advantages:

- Neural networks give amazing flexibility to fit complex and non-linear functions
- For this reason, Neural networks can produce impressive predictive results that exceed any of the algorithms we have previously learned

Meta-Evaluation

Advantages:

- Neural networks give amazing flexibility to fit complex and non-linear functions
- For this reason, Neural networks can produce impressive predictive results that exceed any of the algorithms we have previously learned

Disadvantages:

- Neural networks can be computationally expensive, requiring a lot of time / processing power
- Neural networks function as a “black box” with low interpretive value

Agenda

- Introduction to the Brain
- Introduction to Neural Networks
- Gradient Descent
- Resources for Further Study

Resources

Michael Neilson has a free book called Neural Networks and Deep Learning

- Ch. 1: Using Neural Networks to recognize handwritten digits
- Ch. 2: How the backpropagation algorithm works
- Ch. 3: Improving the way the neural networks learn
- Ch. 4: A visual proof that Neural Networks can compute any function

Resources

Geoffrey Hinton, one of the pioneers of the deep learning movement, has an entire class on Coursera called Neural Networks for Machine Learning

Some of the topics that we didn't cover

- Neural Networks with Feedback Loops – Recursive Neural Networks
- Ensembling Neural Networks – Bayesian Model Averaging
- Unsupervised Neural Networks – Restricted Boltzman Machines (RBM)