Artificial Neural Networks

What if there is no known method for computing the answer?

- Model a complex chemical reaction where precise interactions are not known
- Classify credit card applications
- Forecast stock market
- Recognize hand written characters
- Identify faces, car plate numbers

Learning: Machine or Human

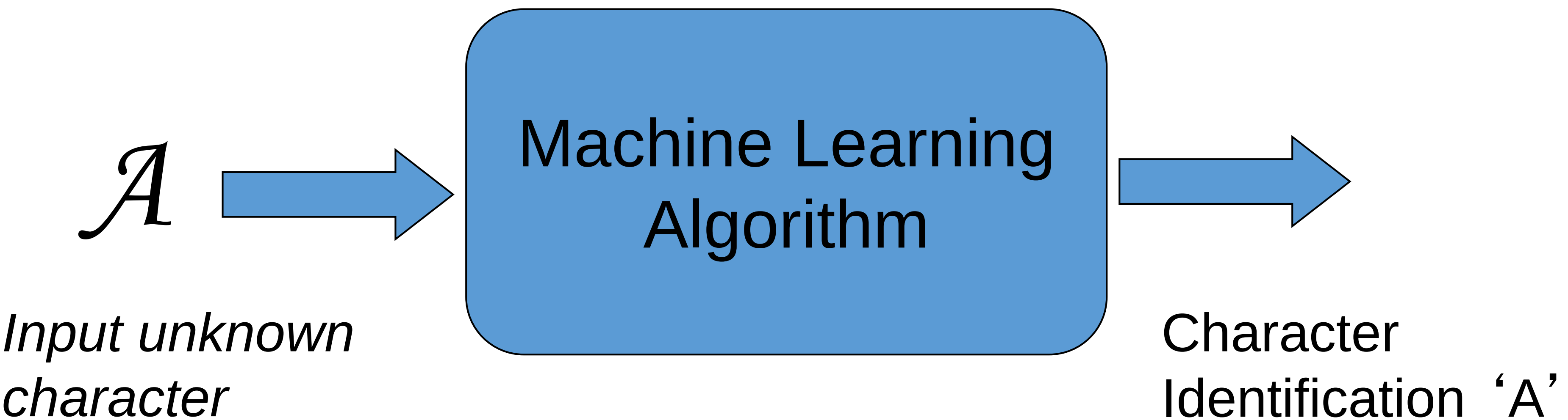
- If there is no known traditional method to solve the problem we can attempt to learn the input - output relationship between the data
  - Statistical methods
  - Methods inspired by ‘biology’

Character Recognition

A A A A A A A A

- *Cannot write a mathematical equation that will recognize each character as an 'A'*
- *But, we can develop a machine learning algorithm that can*

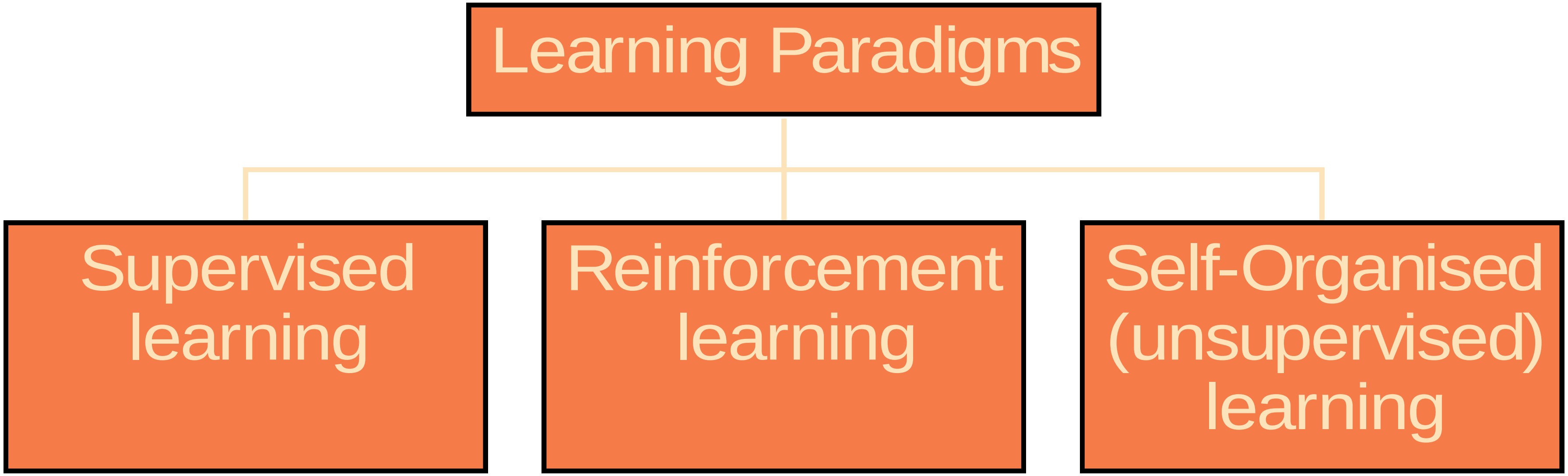
Character Recognition



Artificial Neural Networks

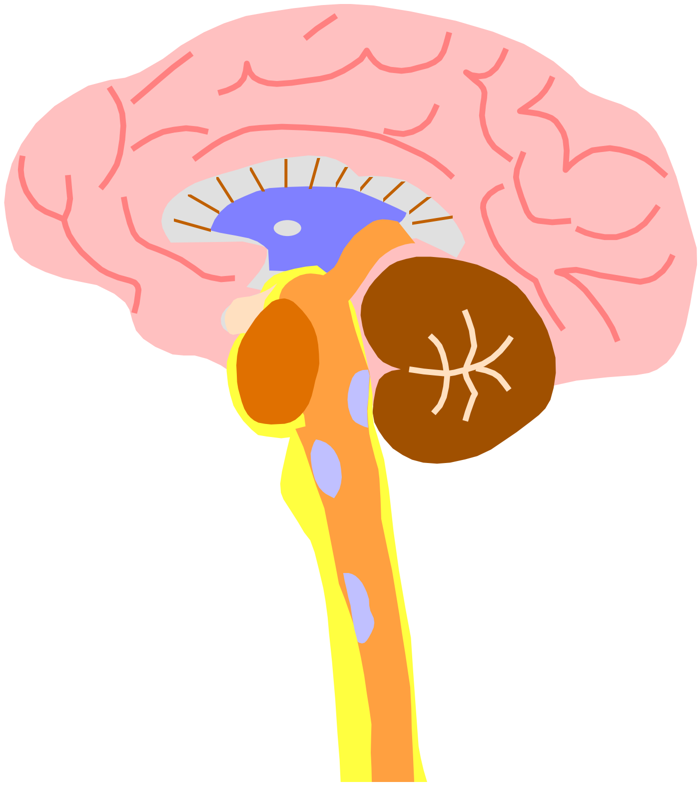
Supervised Framework

ANN Paradigms



Inspiration

- Brain is an *information-processing system*
- Brain is dominant over computers
- Brain can LEARN
- What is the secret of brain's success?



Inspiration

- Is it the speed?
  - NO!

Neurons in brain operate at 10^{-3} sec.

Silicon logic gates operate at 10^{-9} sec.

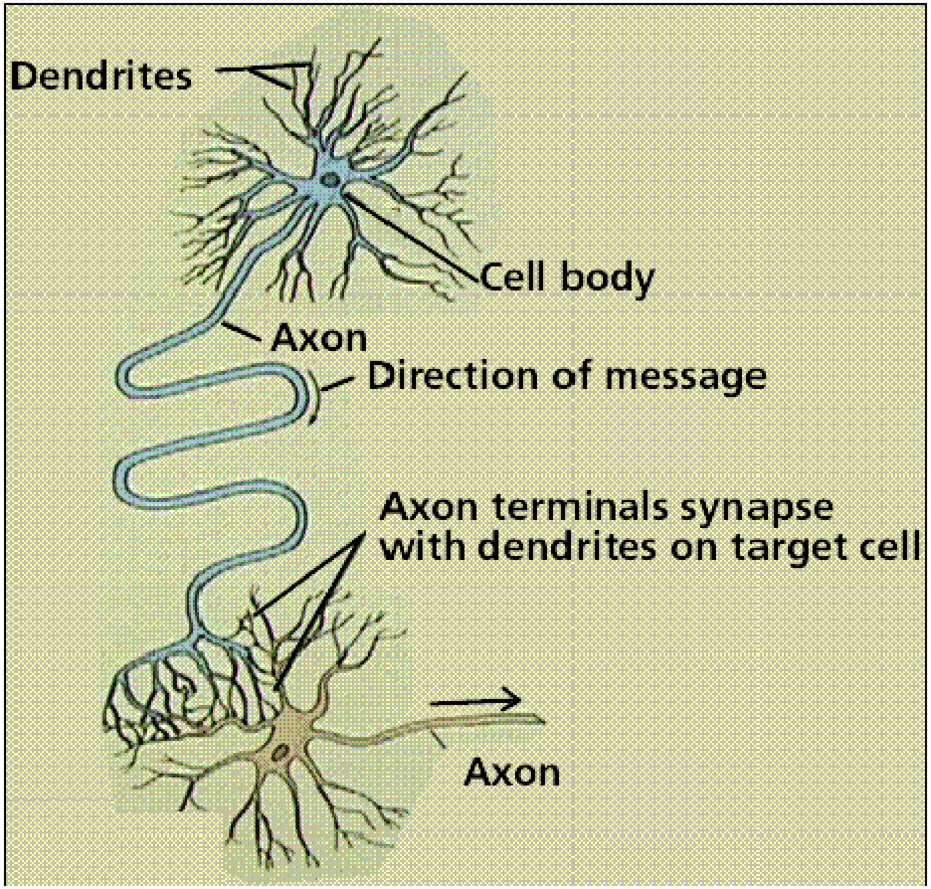
- The trick is in large number of neurons that operate in parallel

There are ca 10^{9} neurons with ca 10^{13} connections between them

Microprocessor speeds vs. neuron speeds

- Current clock speeds exceed 1,800,000,000 clock cycles per second. (1.8 GHz)
- Neurons transfer data at the rate of approximately 100 Hz
- A ‘typical’ computer has one CPU
- The human brain has approximately 10- 100 billions neurons that are highly interconnected
- Each neuron may be connected to 10,000 other neurons

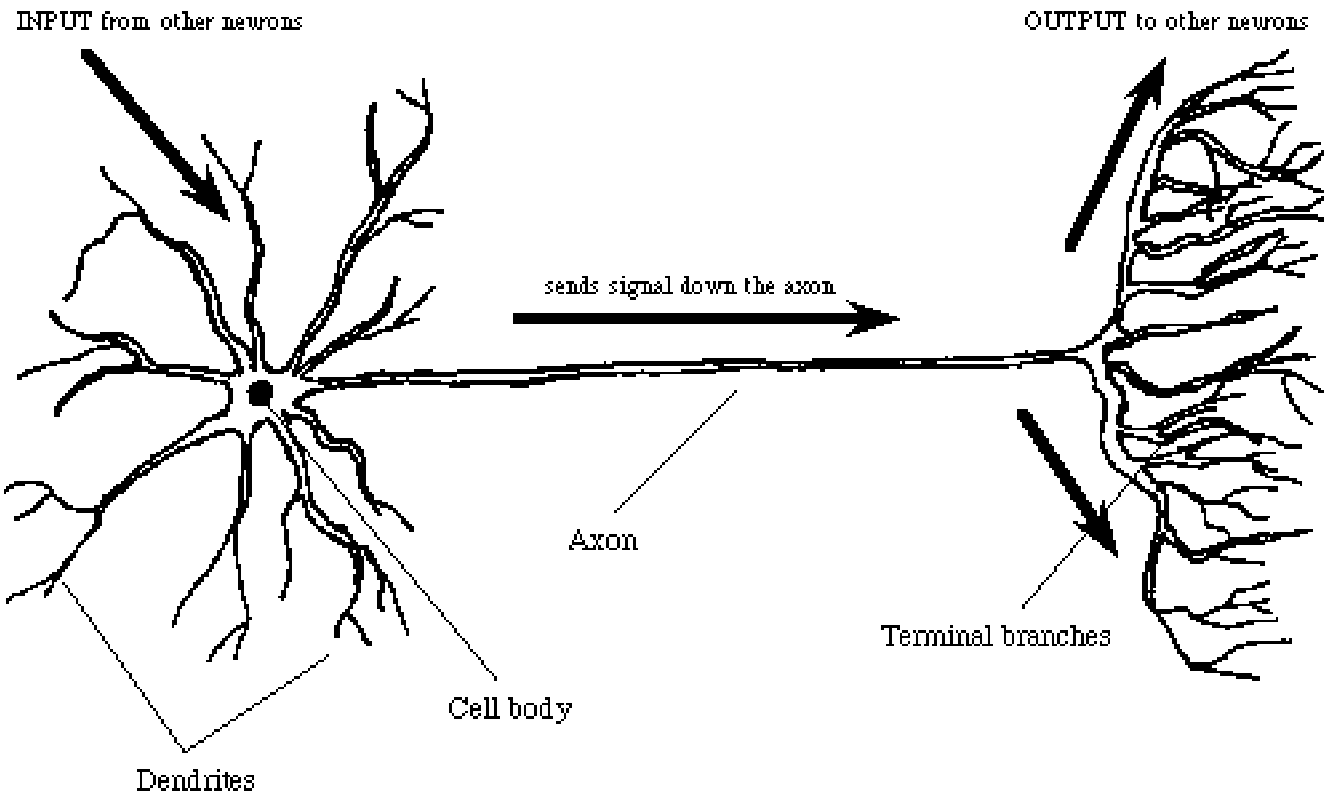
Direction of Nerve Message Transmission



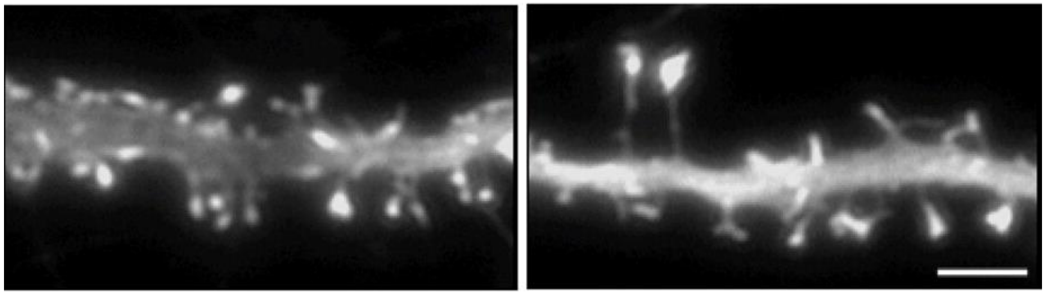
Integrate and Fire Neuron Model

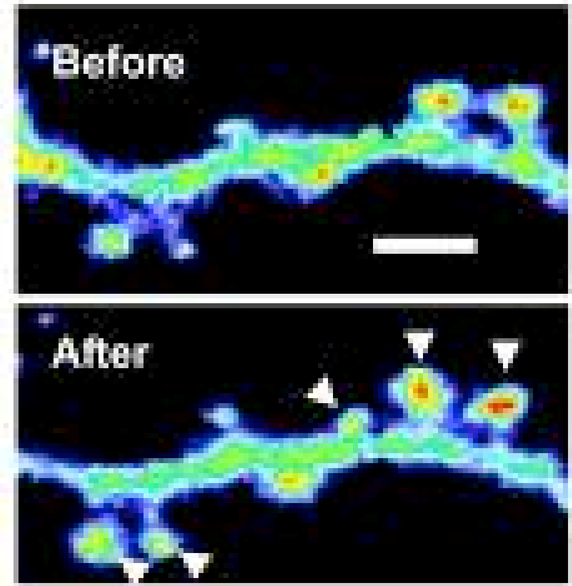
- Neuron receives inputs via synapses on its dendrites
- Each input is weighted based as a function of its synaptic strengths (some inputs more important than others)
- Inputs are integrated together resulting in a change in membrane potential at cell body
- When integrated input exceeds a threshold, the neuron fires a ‘pulse’ down its axon which forms synapses on other neuron’s dendrites

Neurons



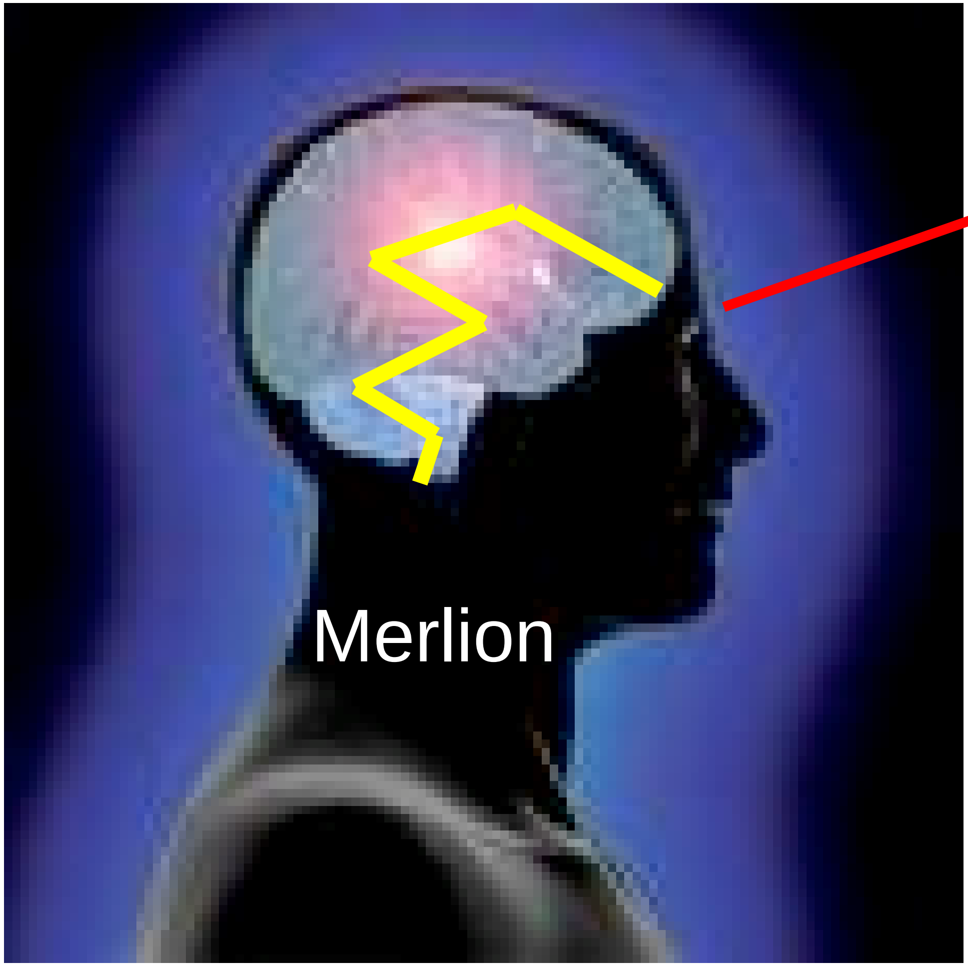
INPUT from other neurons
OUTPUT to other neurons
sends signal down the axon
Axon
Terminal branches
Cell body
Dendrites





- Photographs showing synaptic strengthening before and after stimulus

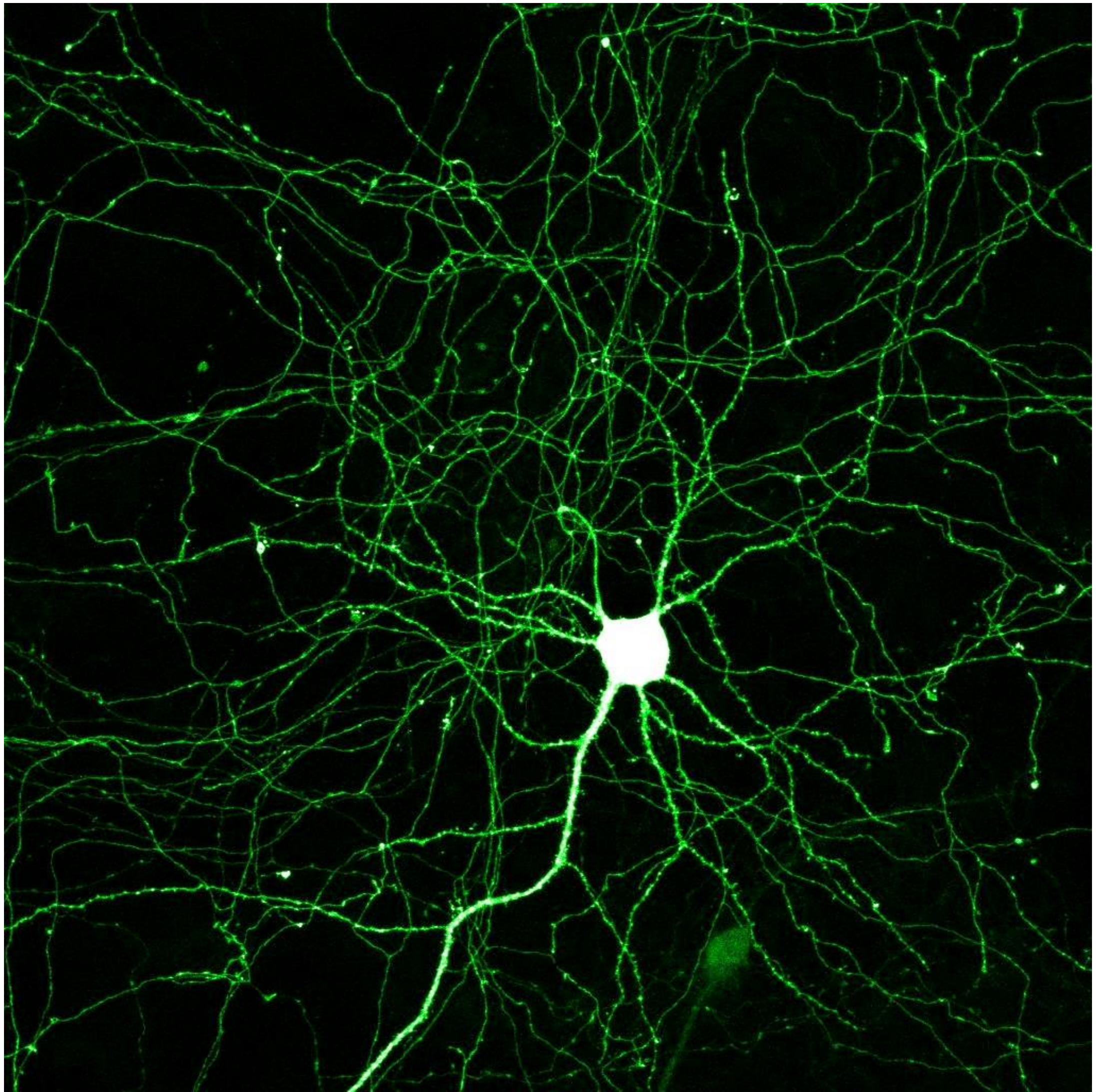
Learning



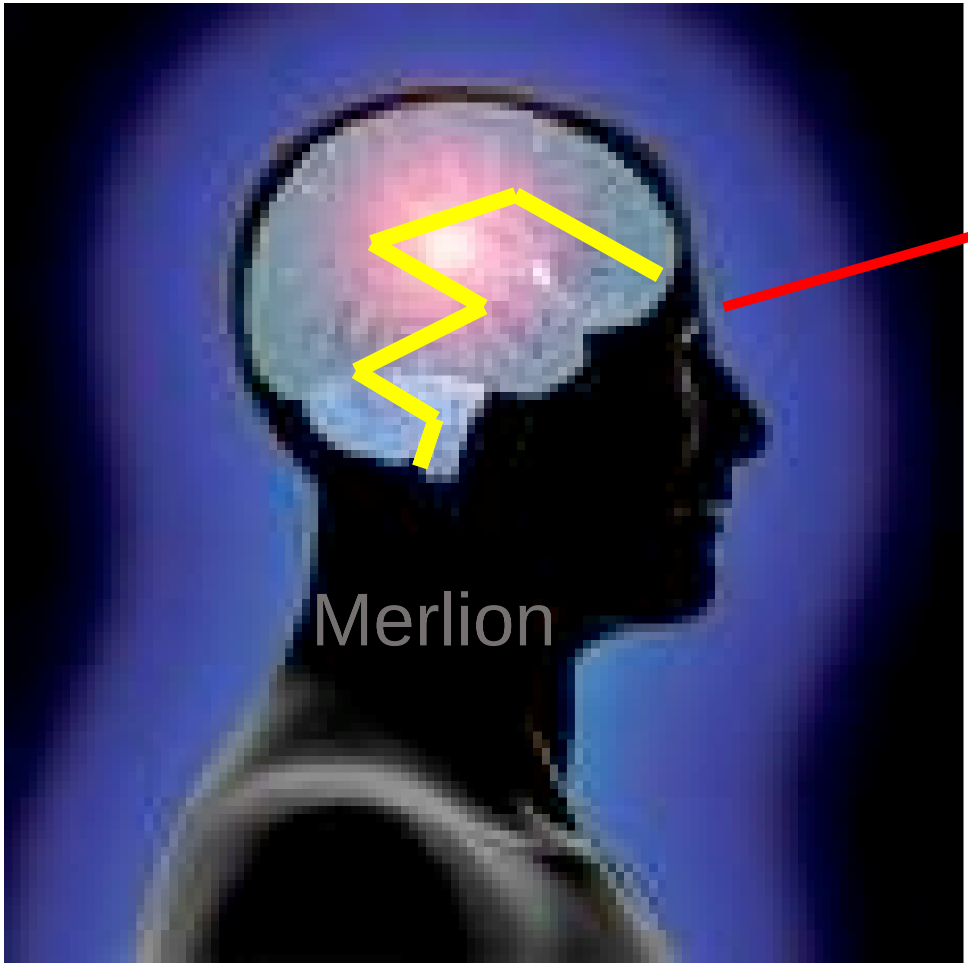
Your brain categorises
Inputs, sounds, videos

Your brain quickly sorts
previous experiences

Brain Scaffold



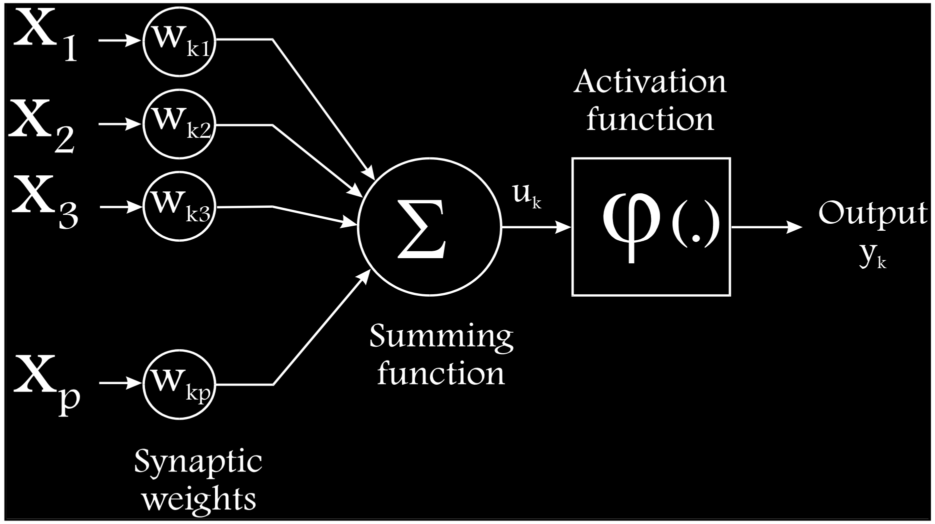
Learning



How can we transform the knowledge and understanding we gain to software and applications?

- Trade-off between neuroscience knowledge and applications
- How and should we develop a model of a brain
  - The brain has the ability to adapt to complex and novel environments
  - Biological models may inspire new computer designs and applications

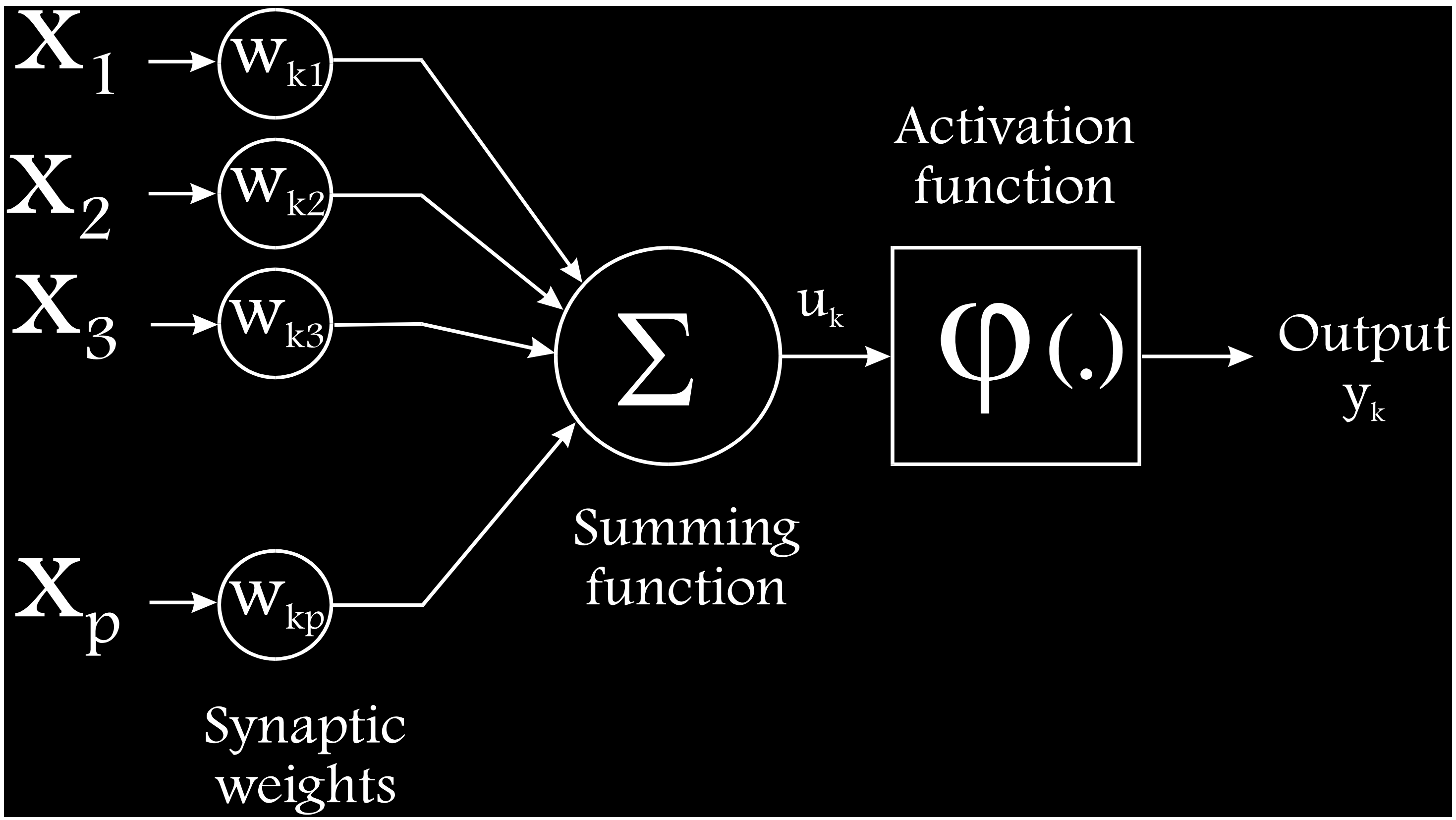
Natural and Artificial Neurons



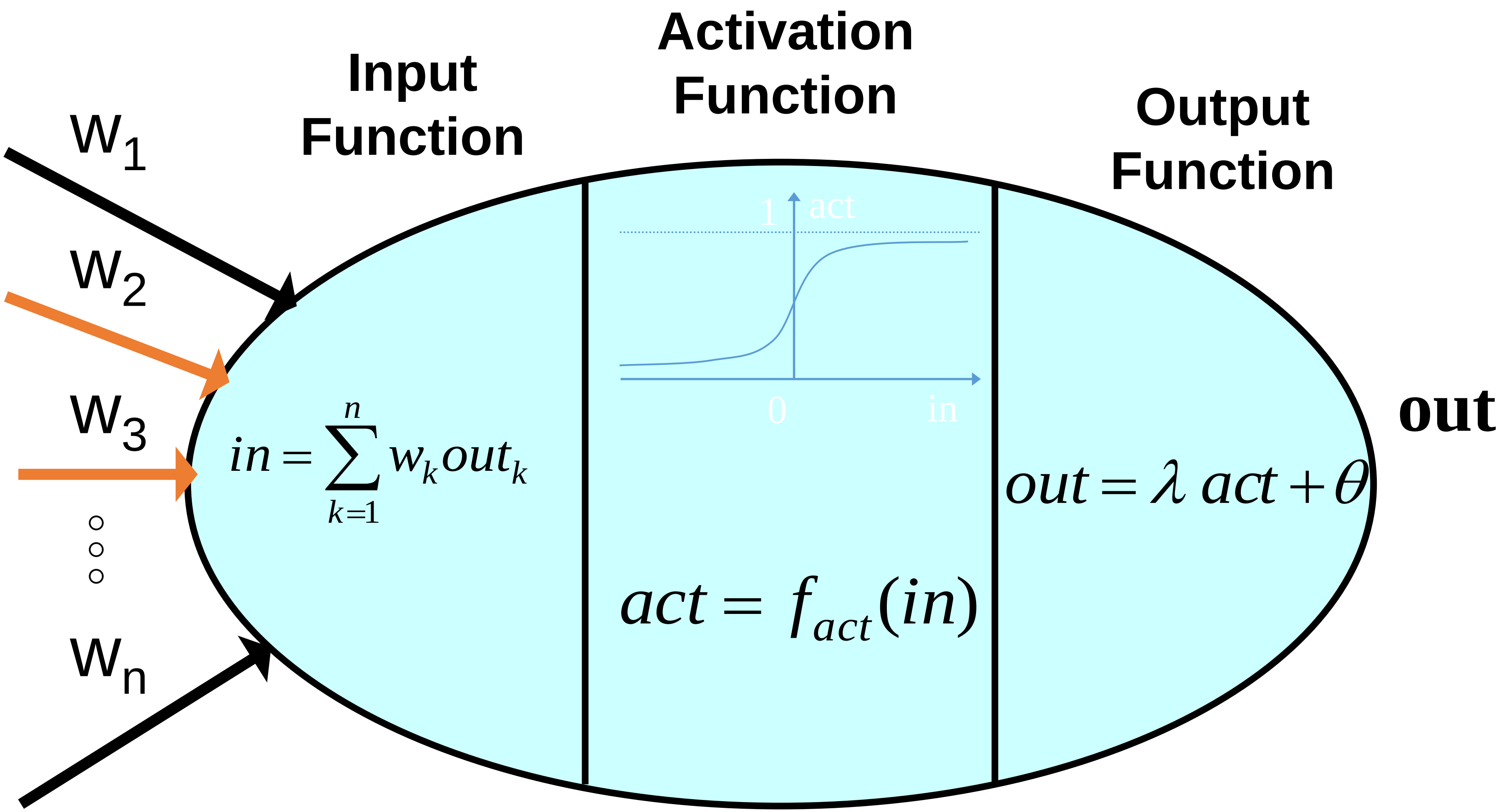
Natural Neuron

Artificial Neuron

Artificial Neuron



Example of a Neuron



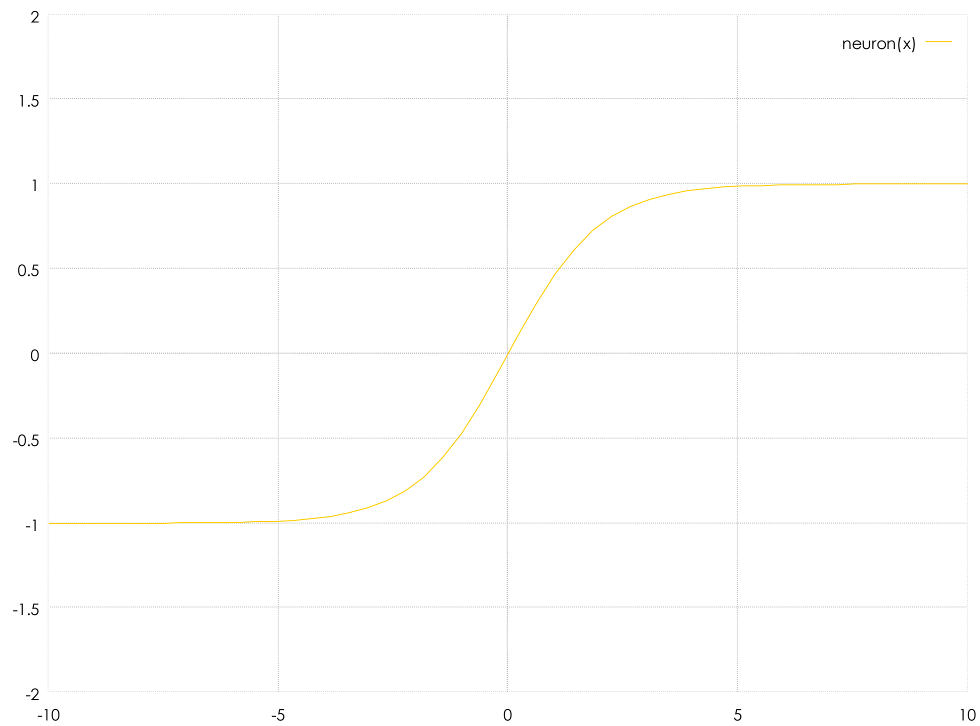
Activation Function

- Linear function

$$y = \sum_{i=1}^{r} w_i x_i$$

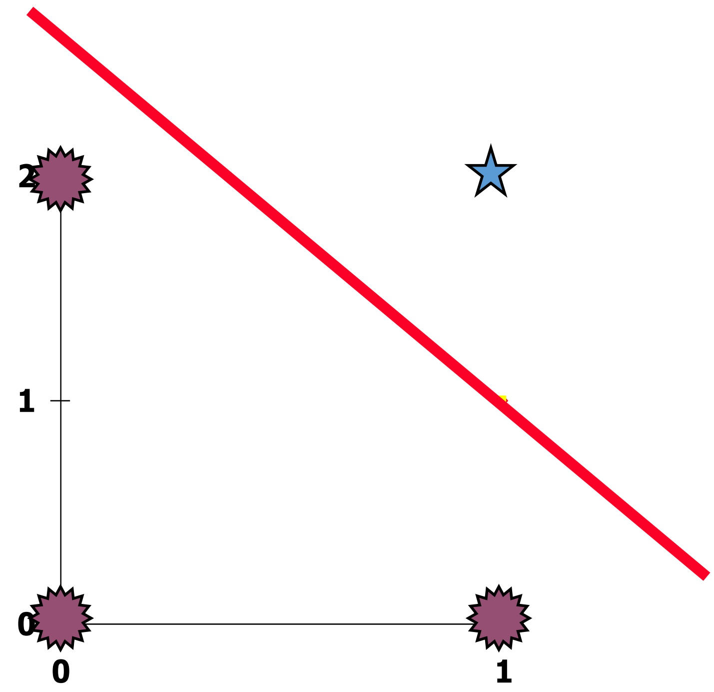
- Sigmoid

$$y = \frac{1 + e^{(\sum_i w_i x_i + \theta)}}{1 + e^{-(\sum_i w_i x_i + \theta)}}$$



What can single neuron do?

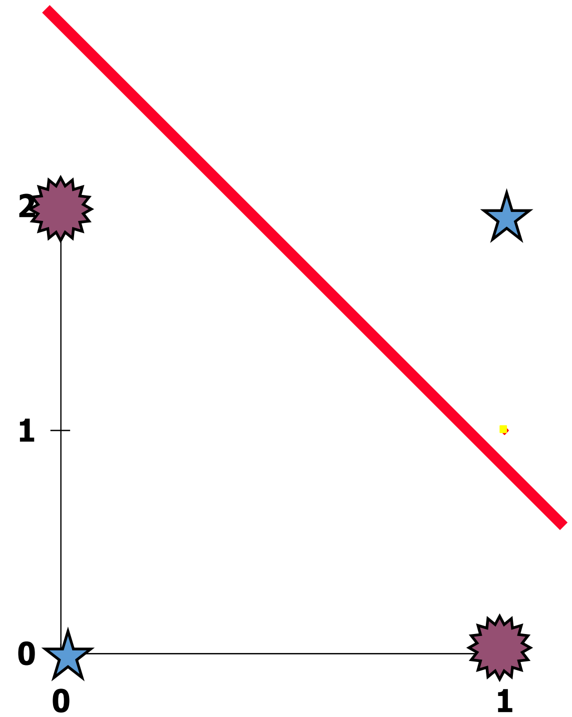
- Not much!!
- Single neuron can be trained to correctly solve so-called linearly-separable problem (such as OR)



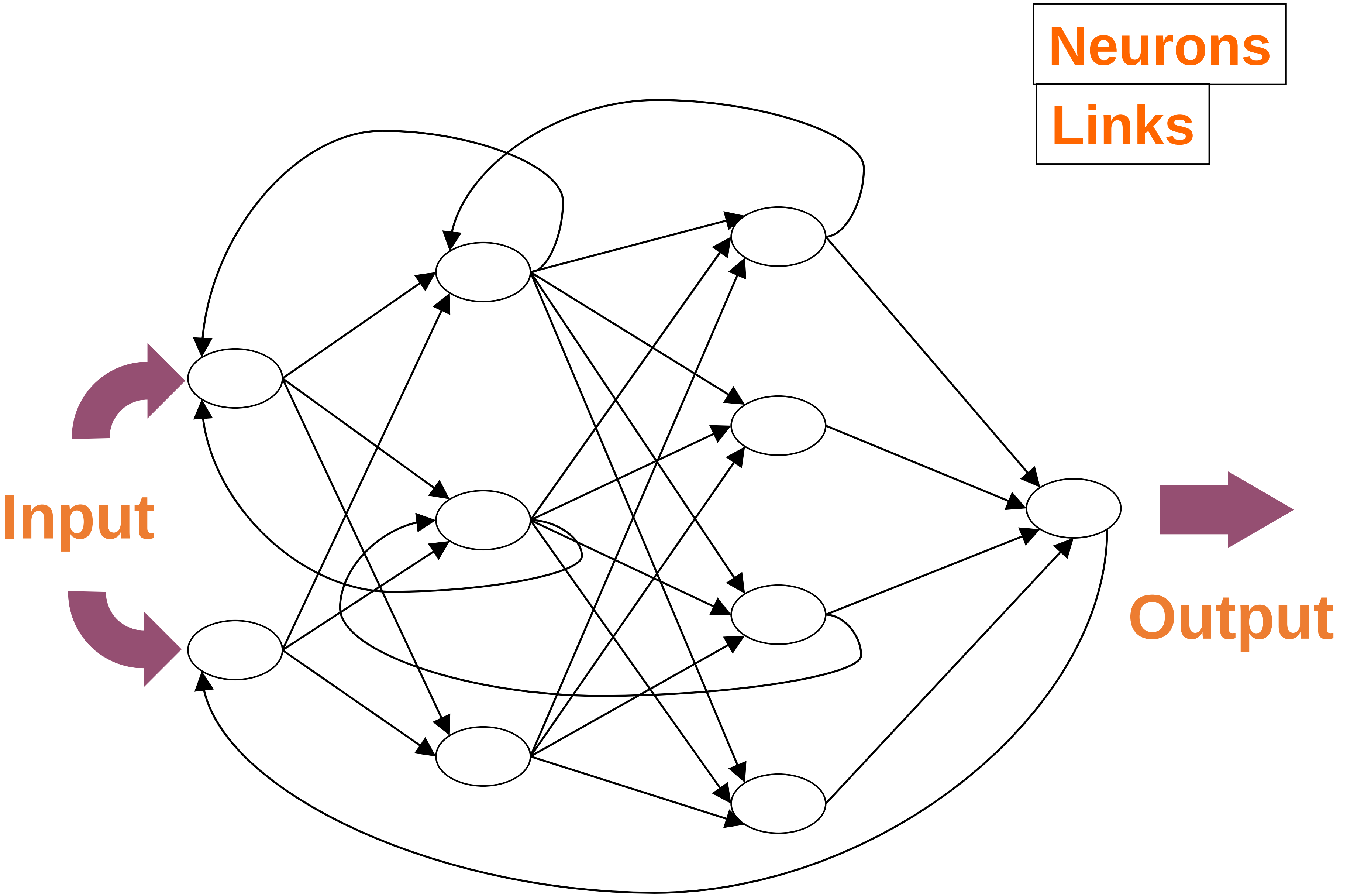
Input 1	Input 2	OR
1	1	1
1	0	1
0	1	1
0	0	0

What can single neuron do?

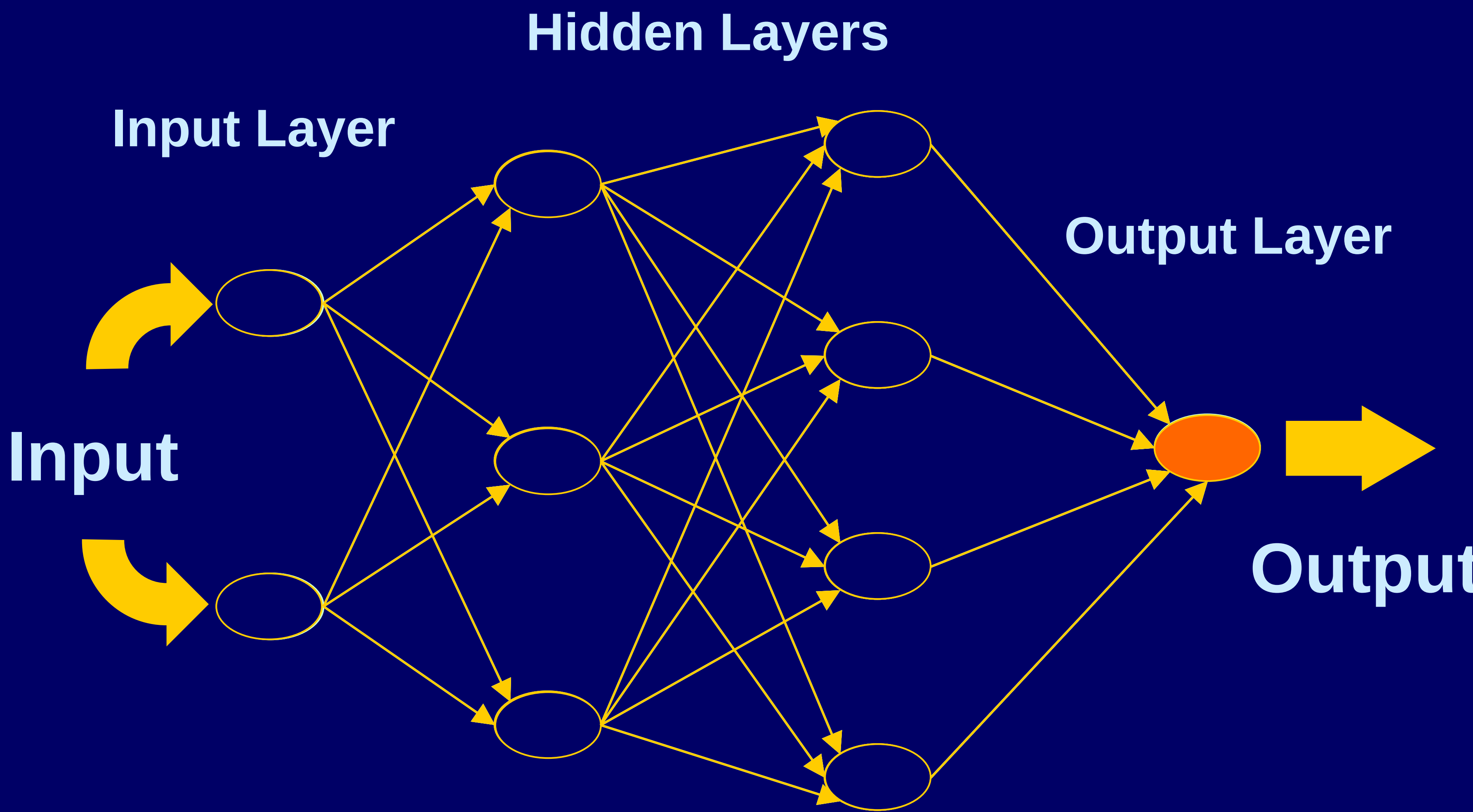
- More complicated problems (non-linear) **cannot** be solved
- i.e. XOR
- For these and more difficult problems, people applied networks of interconnected neurons - thus ANNs



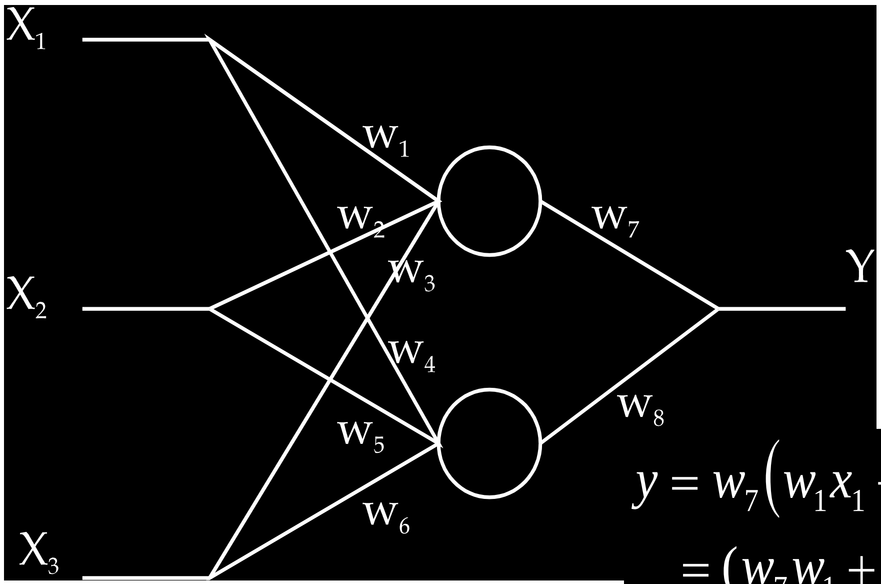
Artificial Neural Network



The Propagation of a Signal in a Feedforward NN



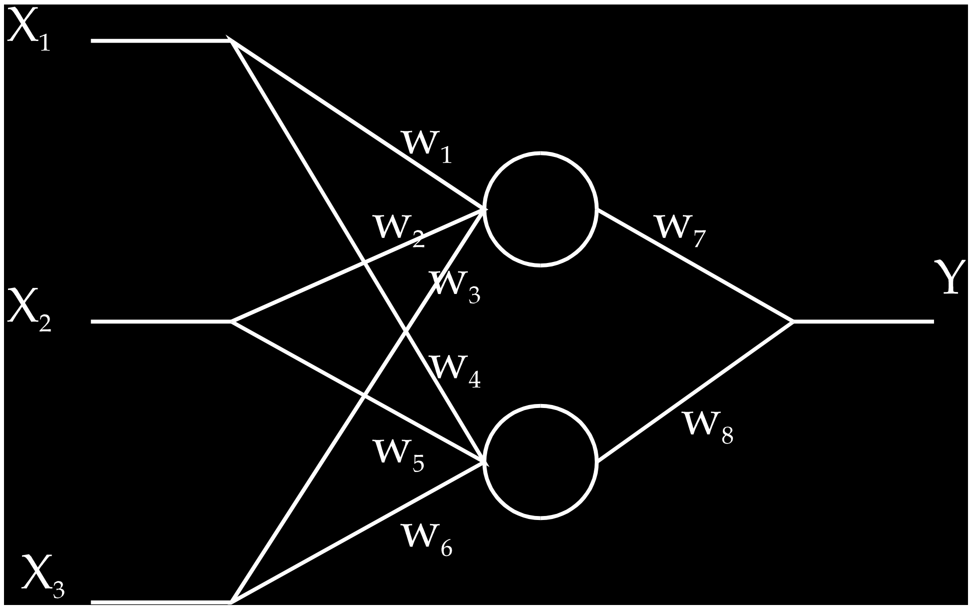
What Does a Network of Linear Neurons Correspond to?



$$
\begin{aligned}
y &= w_7\left(w_1x_1 + w_2x_2 + w_3x_3\right) + w_8\left(w_4x_1 + w_5x_2 + w_6x_3\right) \\
&= (w_7w_1 + w_8w_4)x_1 + (w_7w_2 + w_8w_5)x_2 + (w_7w_3 + w_8w_6)x_3 \\
&= \overline{w}_1x_1 + \overline{w}_2x_2 + \overline{w}_3x_3
\end{aligned}
$$

$$
\begin{aligned}
\overline{w}_1 &= w_7w_1 + w_8w_4 \\
\overline{w}_2 &= w_7w_2 + w_8w_5 \\
\overline{w}_3 &= w_7w_3 + w_8w_6
\end{aligned}
$$

ANN as a Set of Equations



$$y = \frac{1+e^{w_7\tilde{x}_1+w_8\tilde{x}_2+\theta_3}}{1+e^{-w_7\tilde{x}_1-w_8\tilde{x}_2-\theta_3}}$$

$$\tilde{x}_1 = \frac{1+e^{w_1x_1+w_2x_2+w_3x_3+\theta_1}}{1+e^{-w_1x_1-w_2x_2-w_3x_3-\theta_1}}$$

$$\tilde{x}_2 = \frac{1+e^{w_4x_1+w_5x_2+w_6x_3+\theta_2}}{1+e^{-w_4x_1-w_5x_2-w_6x_3-\theta_2}}$$

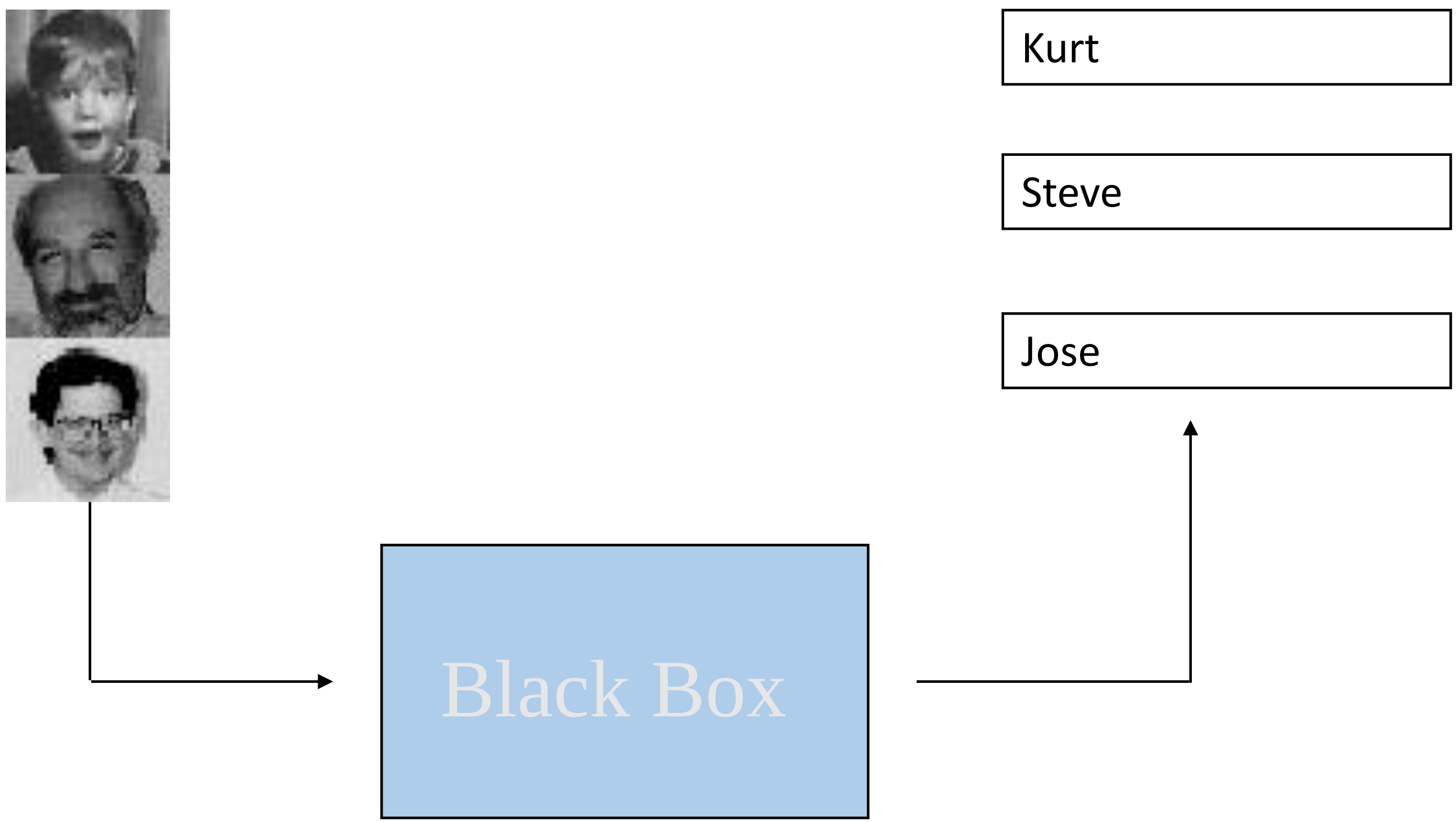
Artificial Neural Network

- Over parameterized set of equations
- Global
- Non-linear

Artificial Neural Networks

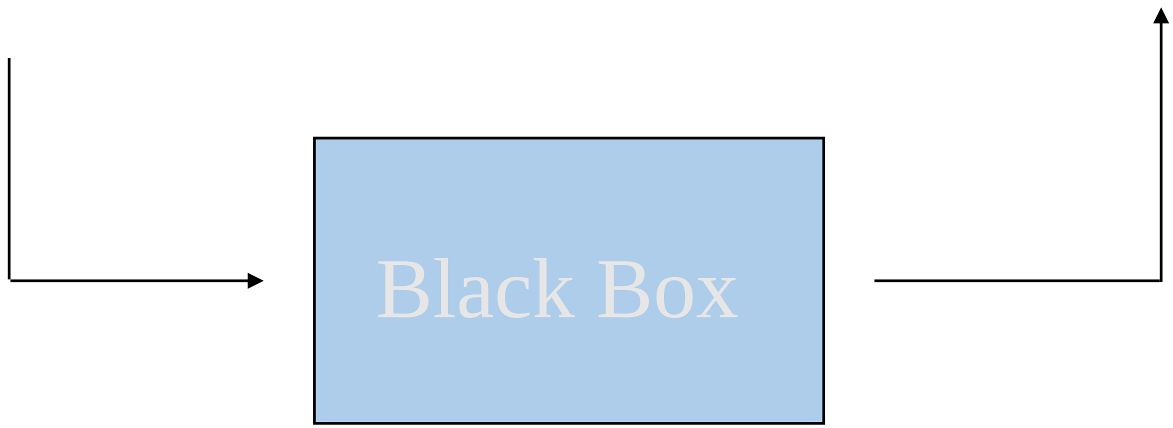
- Incoming signal is transformed to output signal within a processing unit by passing it through activation function
- Network is *trained* to produce known or desired output responses for given input
- Interconnection weights are adjusted during training in a learning algorithm

The ultimate black box



The ultimate black box

1 34 57 255 53 54 87…	23 87 36 98 09…
1 43 54 34 65 45 23…	76 44 82 17 83…
…	…
4 56 76 53 98 45 12…	72 69 52 78 28…



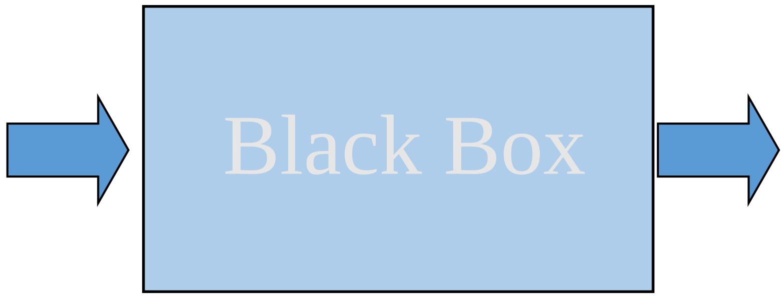
The ultimate black box

1 34 57 255 53 54 87…

1 43 54 34 65 45 23…

…

4 56 76 53 98 45 12…



23 87 36 98 09…

76 44 82 17 83…

…

72 69 52 78 28…

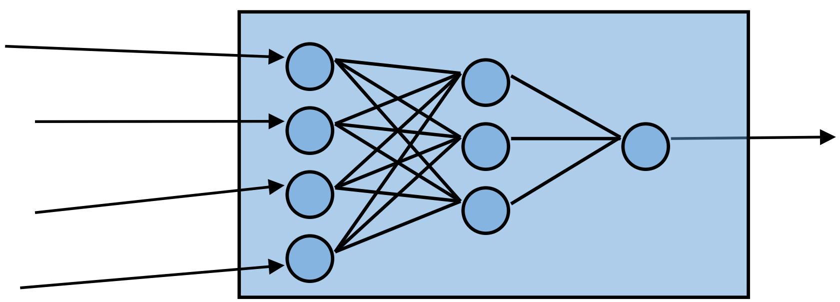
The ultimate black box

1 34 57 255 53 54 87…

1 43 54 34 65 45 23…

…

4 56 76 53 98 45 12…



23 87 36 98 09…

76 44 82 17 83…

…

72 69 52 78 28…

Learning

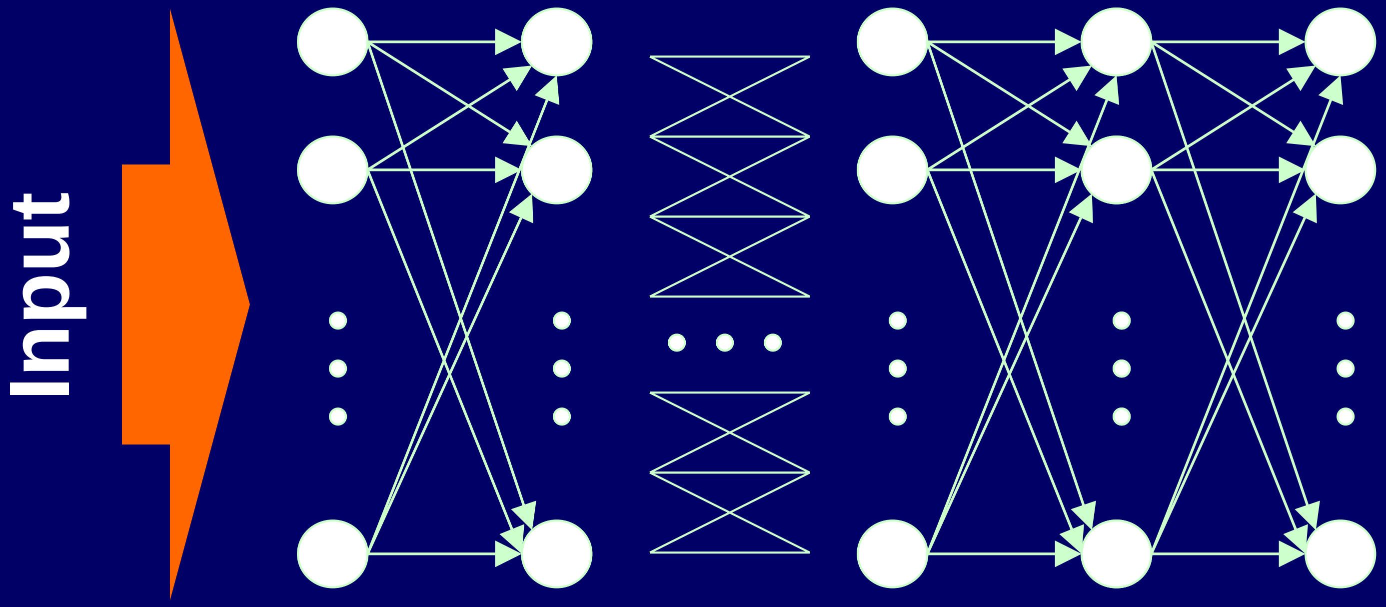
Definition

Learning is a process by which the free parameters of a neural network are adapted through a continuing process of stimulation by the environment in which the network is embedded. The type of learning is determined by the manner in which the parameter changes take place.

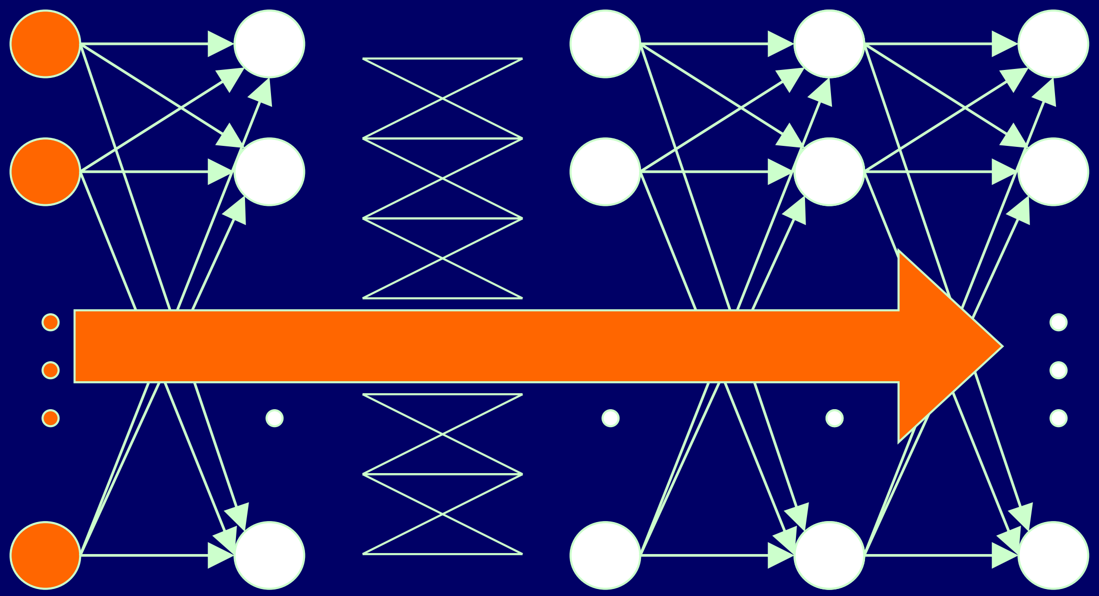
Error-Correction Learning

- Learning → Minimisation of Error
- One can minimise error in many different ways
- Probably the most popular algorithm (for supervised learning) is referred to as **backpropagation of error**, which is a gradient-based search method

Basics of BackPropagation Algorithm

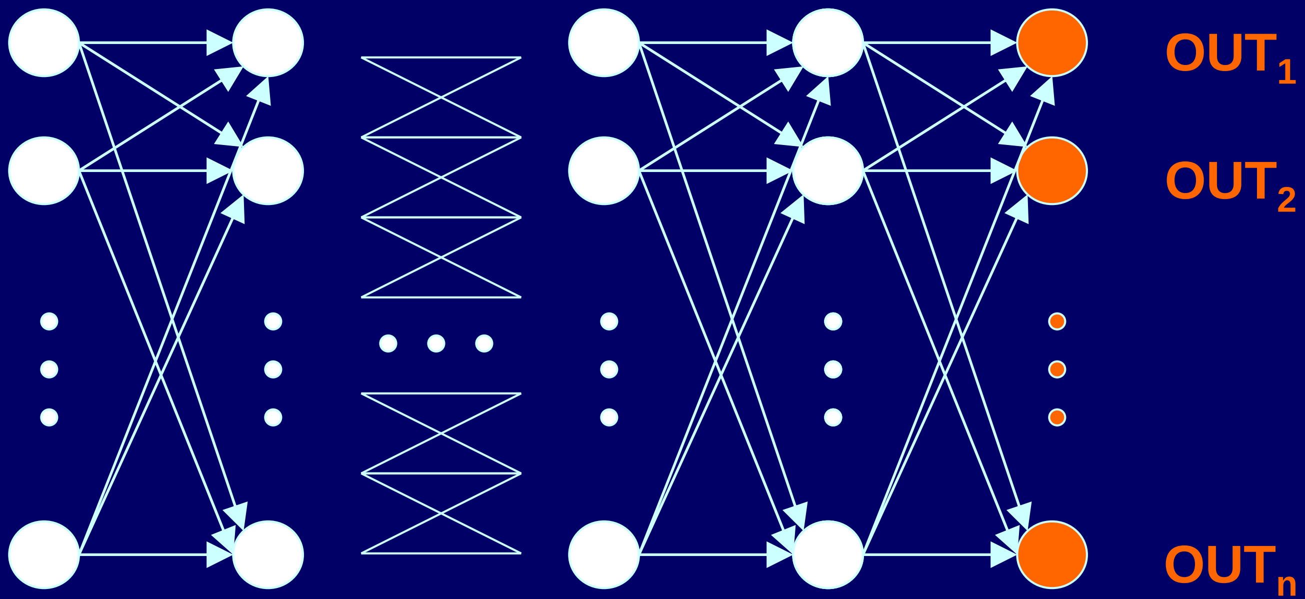


Basics of BackPropagation Algorithm



Basics of BackPropagation Algorithm

$$J = \frac{1}{2}\sum_{i=1}^{n}(desired_i - out_i)^2$$



Basics of the Backpropagation Algorithm

$$\Delta w_{ji} = -\gamma \delta_j(n) out_i(n)$$

IF *j* is an Output Neuron

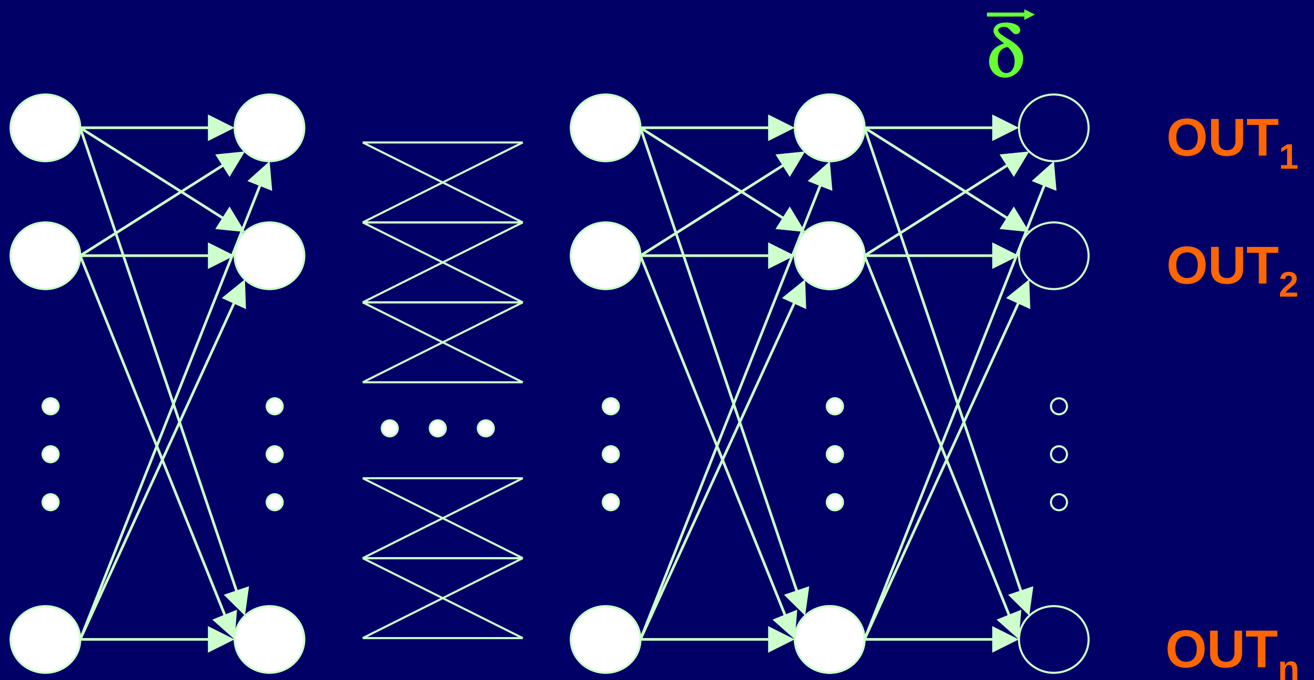
$$\delta_j(n) = \left(desired_j(n) - out_j(n)\right) f_{act}'(in_j(n))$$

IF *j* is a Hidden Neuron

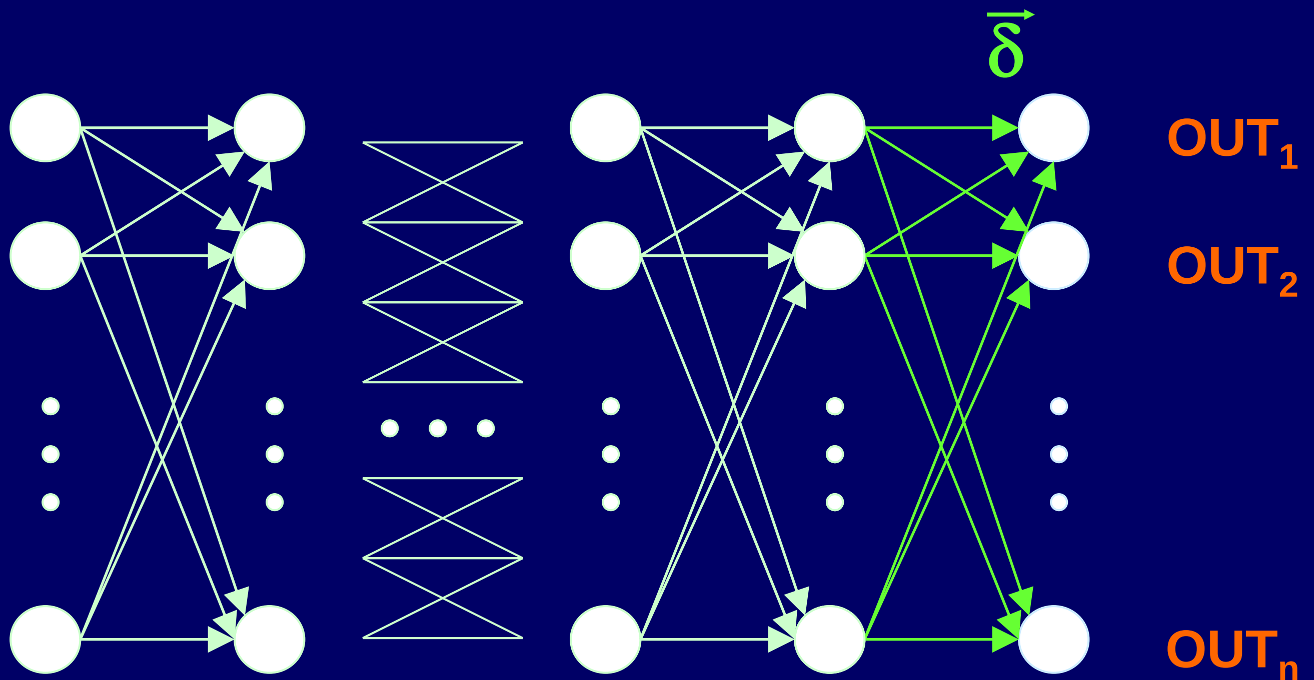
***k* is in the next layer**

$$\delta_j(n) = f_{act}'(in_j(n)) \sum_k \delta_{kj}(n) w_{kj}(n)$$

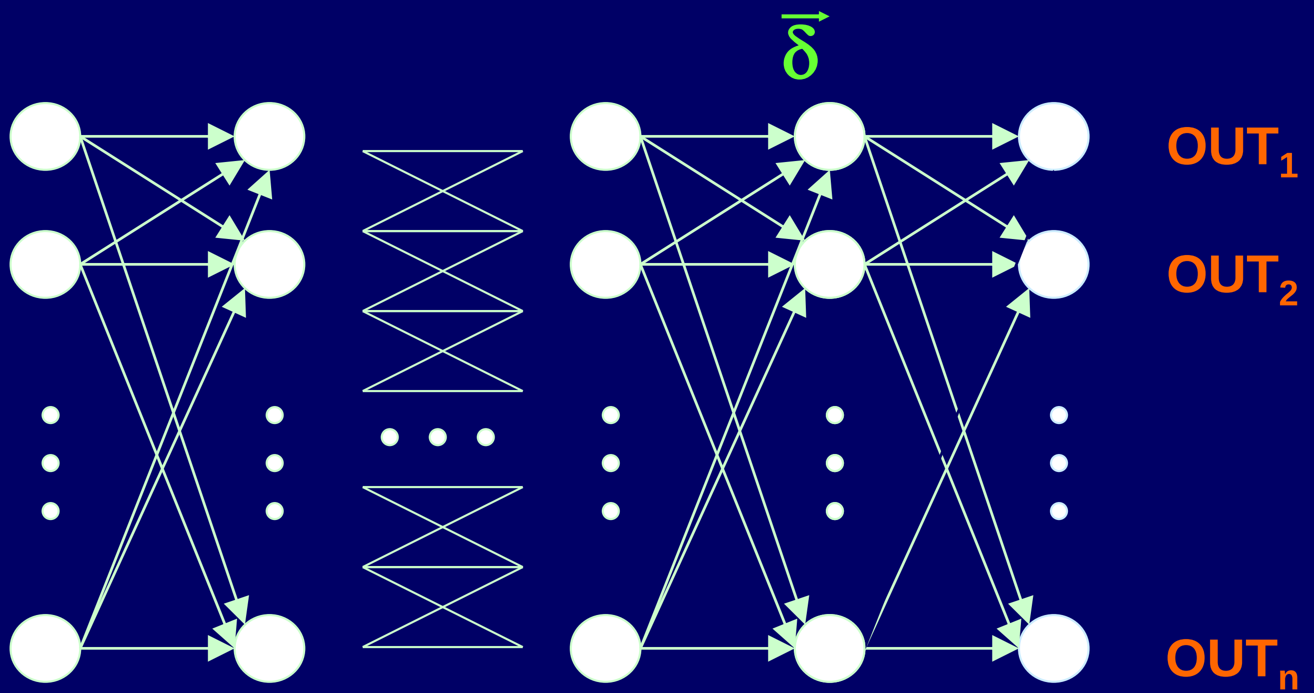
Basics of BackPropagation Algorithm



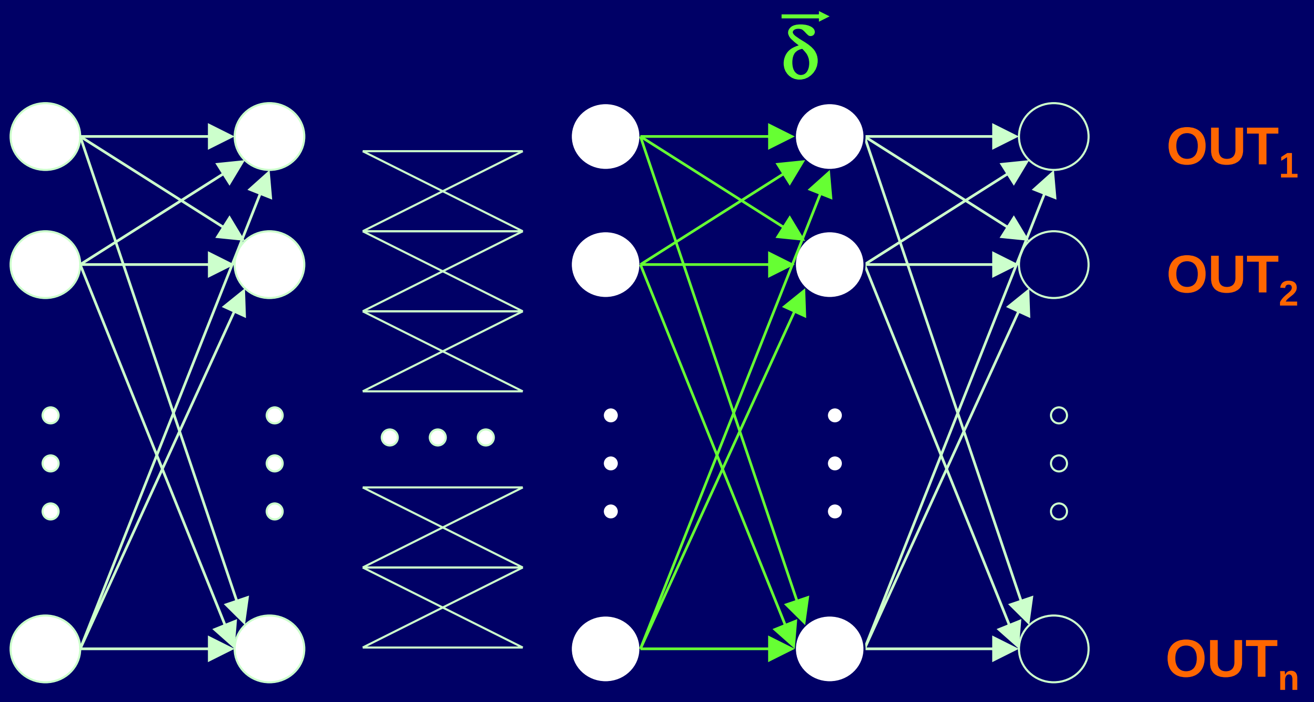
Basics of BackPropagation Algorithm



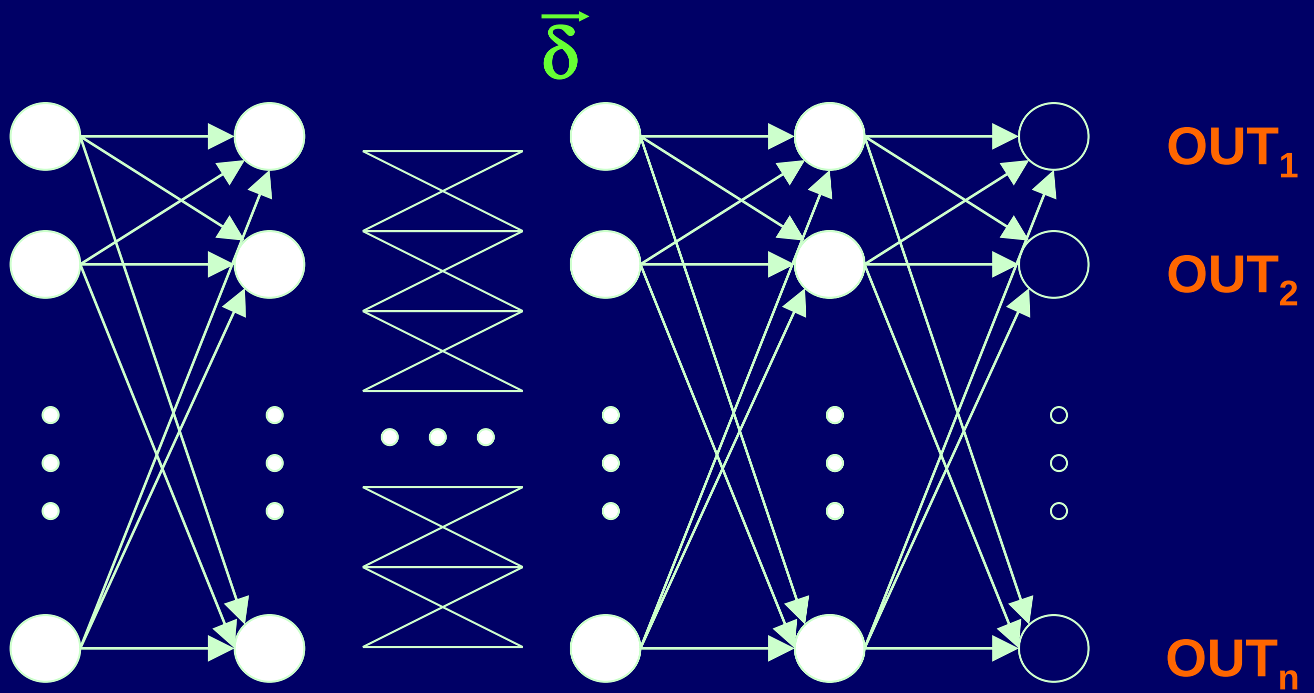
Basics of BackPropagation Algorithm



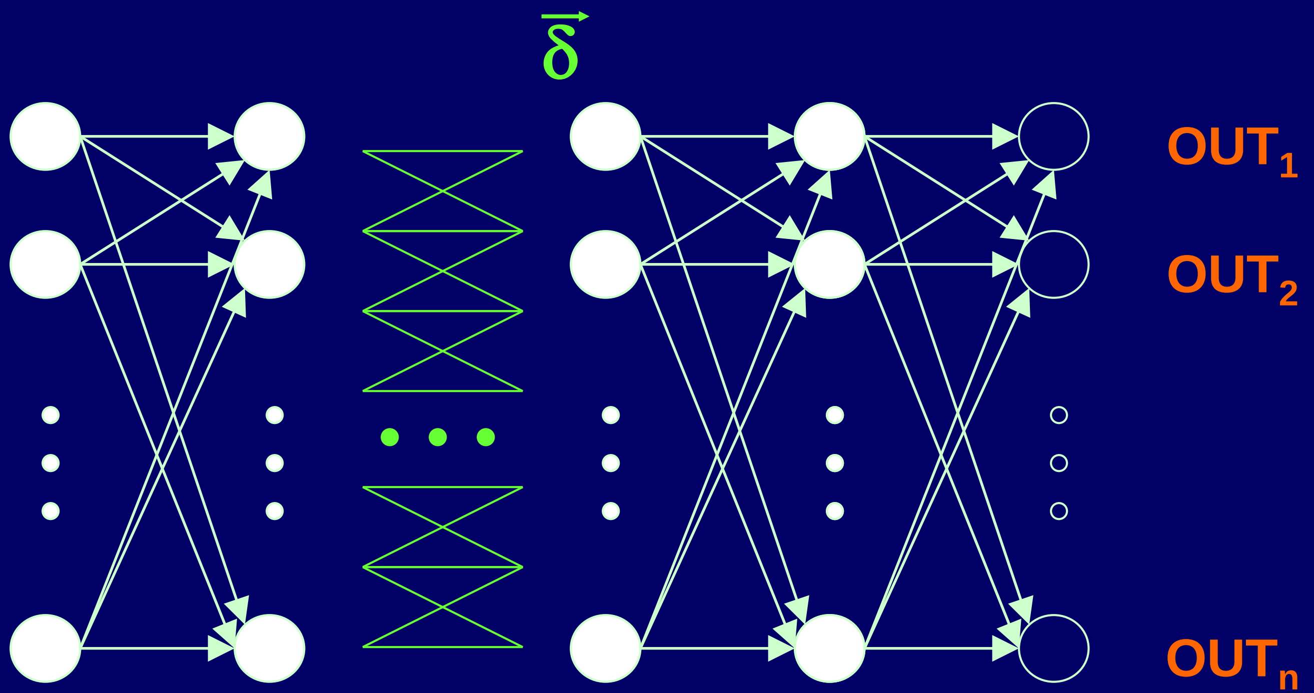
Basics of BackPropagation Algorithm



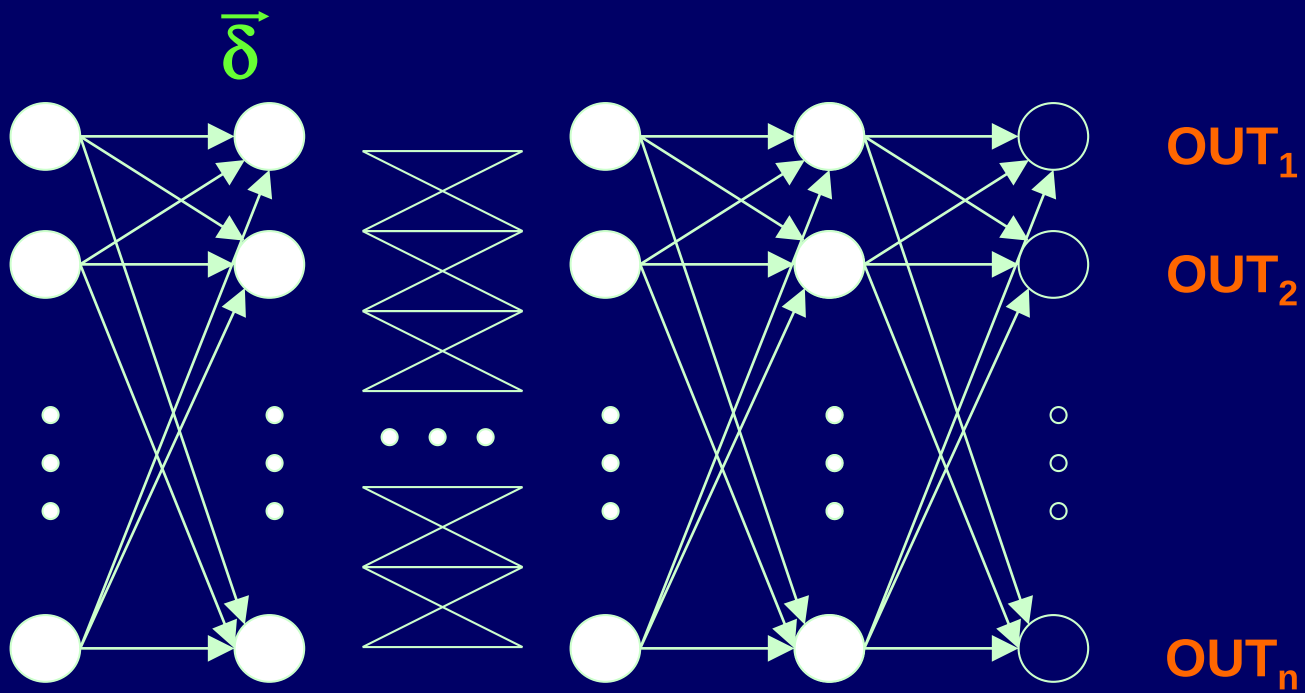
Basics of BackPropagation Algorithm



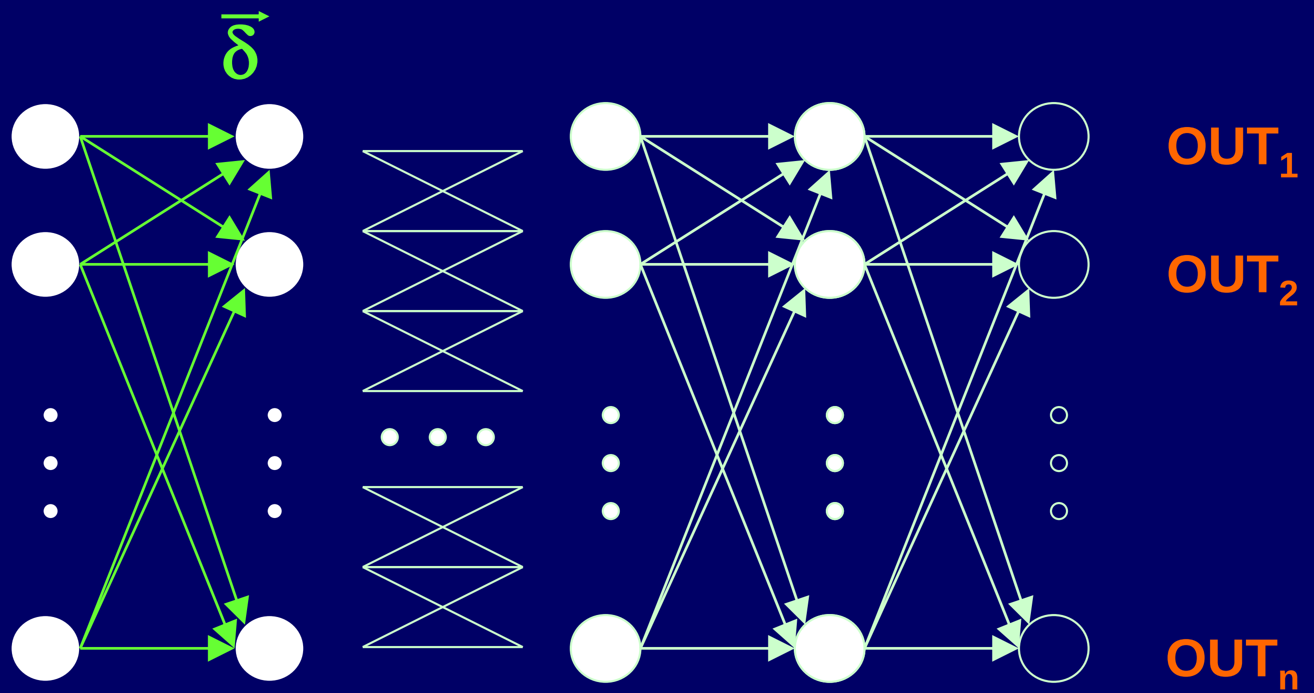
Basics of BackPropagation Algorithm



Basics of BackPropagation Algorithm



Basics of BackPropagation Algorithm



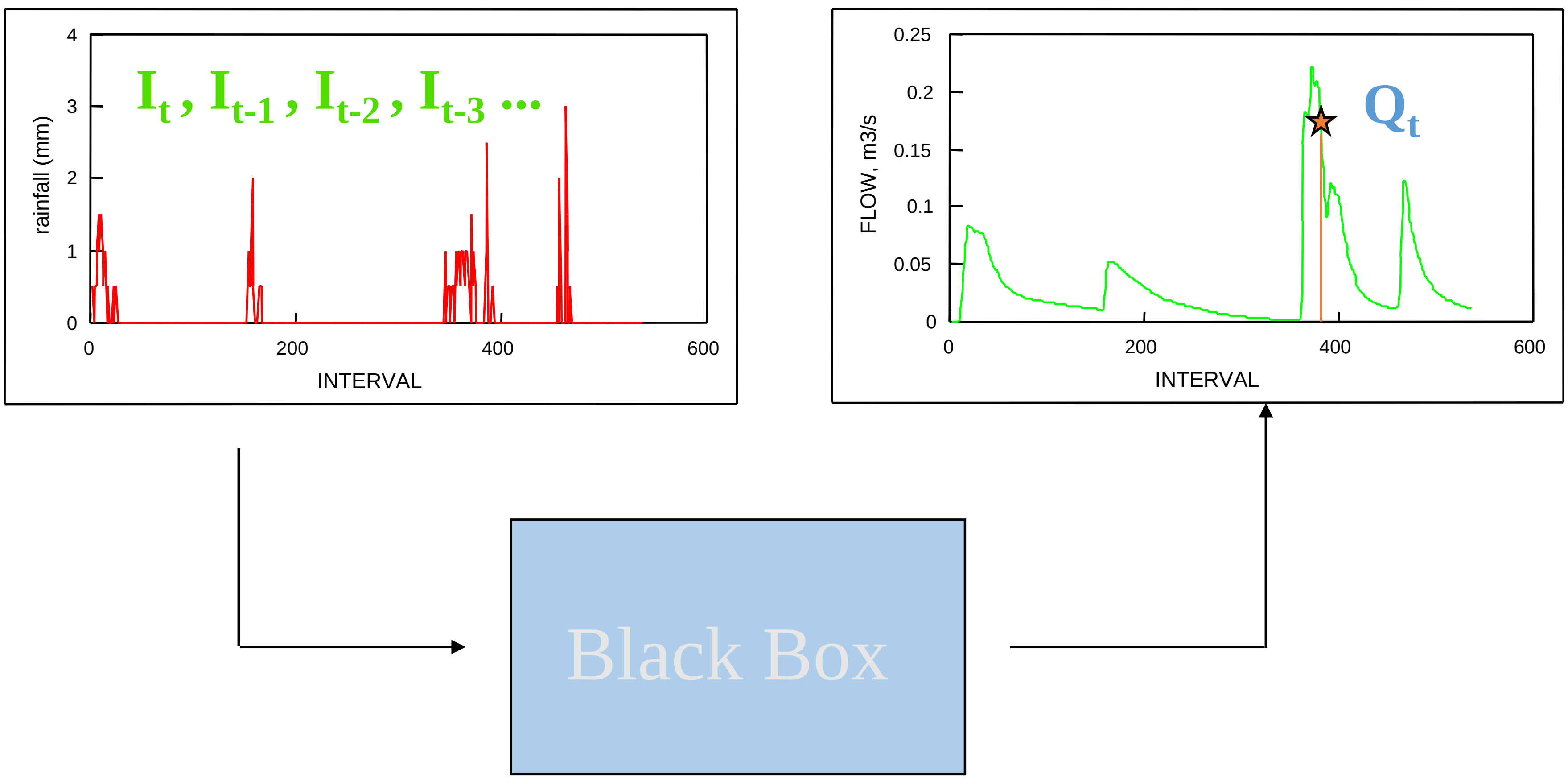
Error-Correction Learning pseudo code

1. Start with random collection of weights
2. Calculate error (*e.g.* J)
3. Correct weights according to some rule (*e.g.* *delta* rule)

$$\Delta w_{kj}(n) = \Delta e_k(n)\, x_j(n)$$

4. Go to step 2

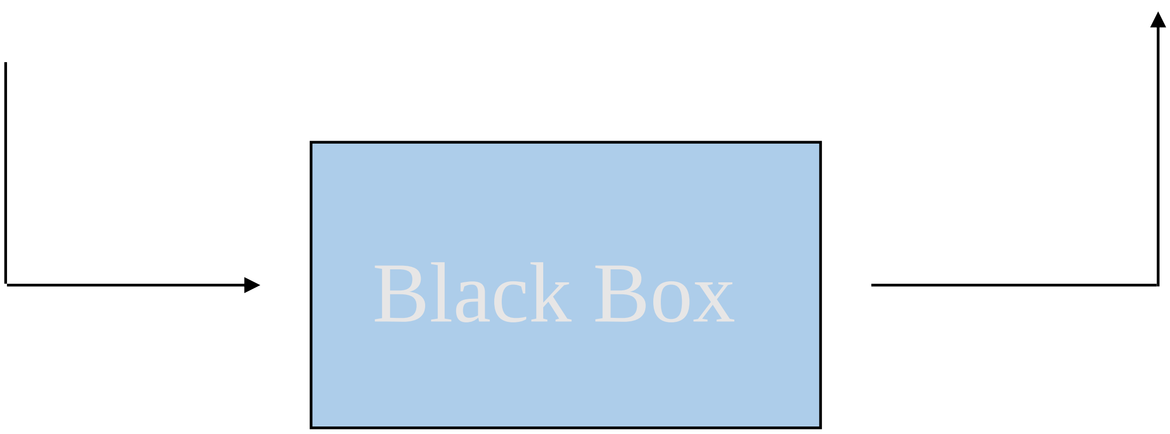
The ultimate black box



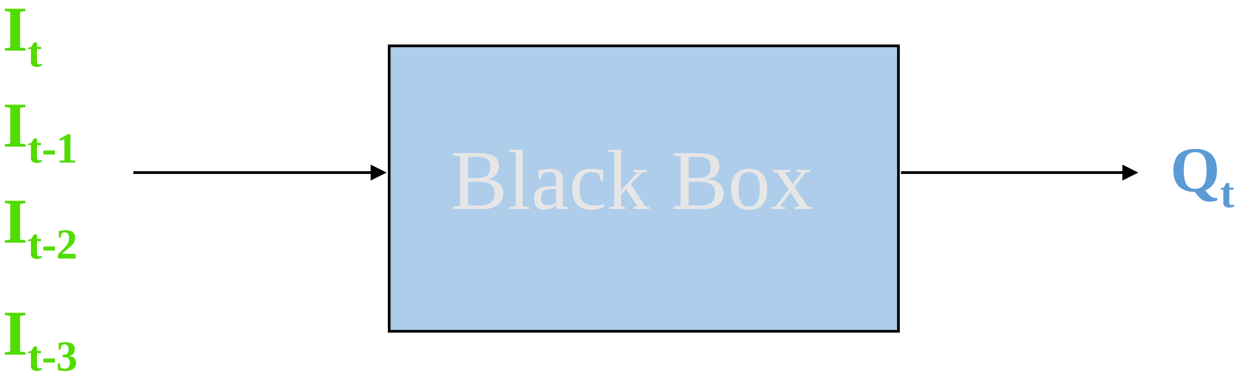
The ultimate black box

$I_t, I_{t-1}, I_{t-2}, I_{t-3} \ldots$

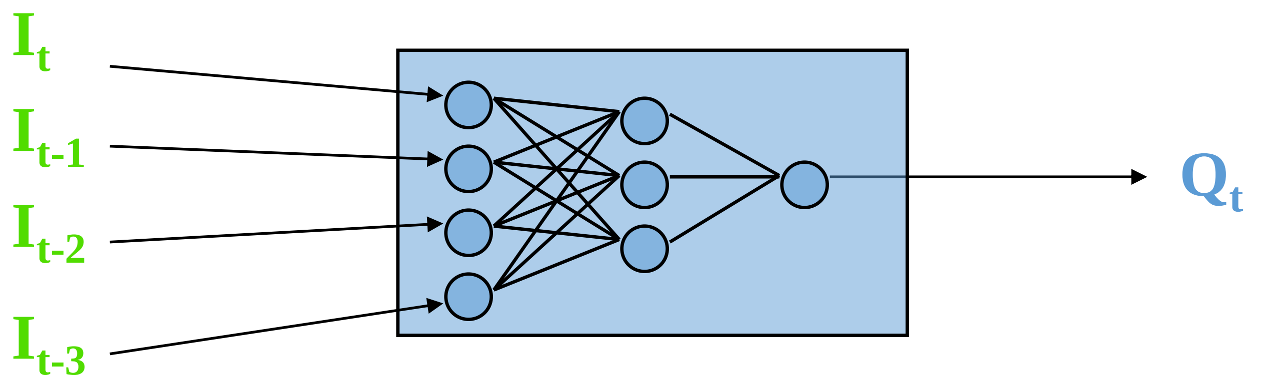
Q_t



The ultimate black box



The ultimate black box

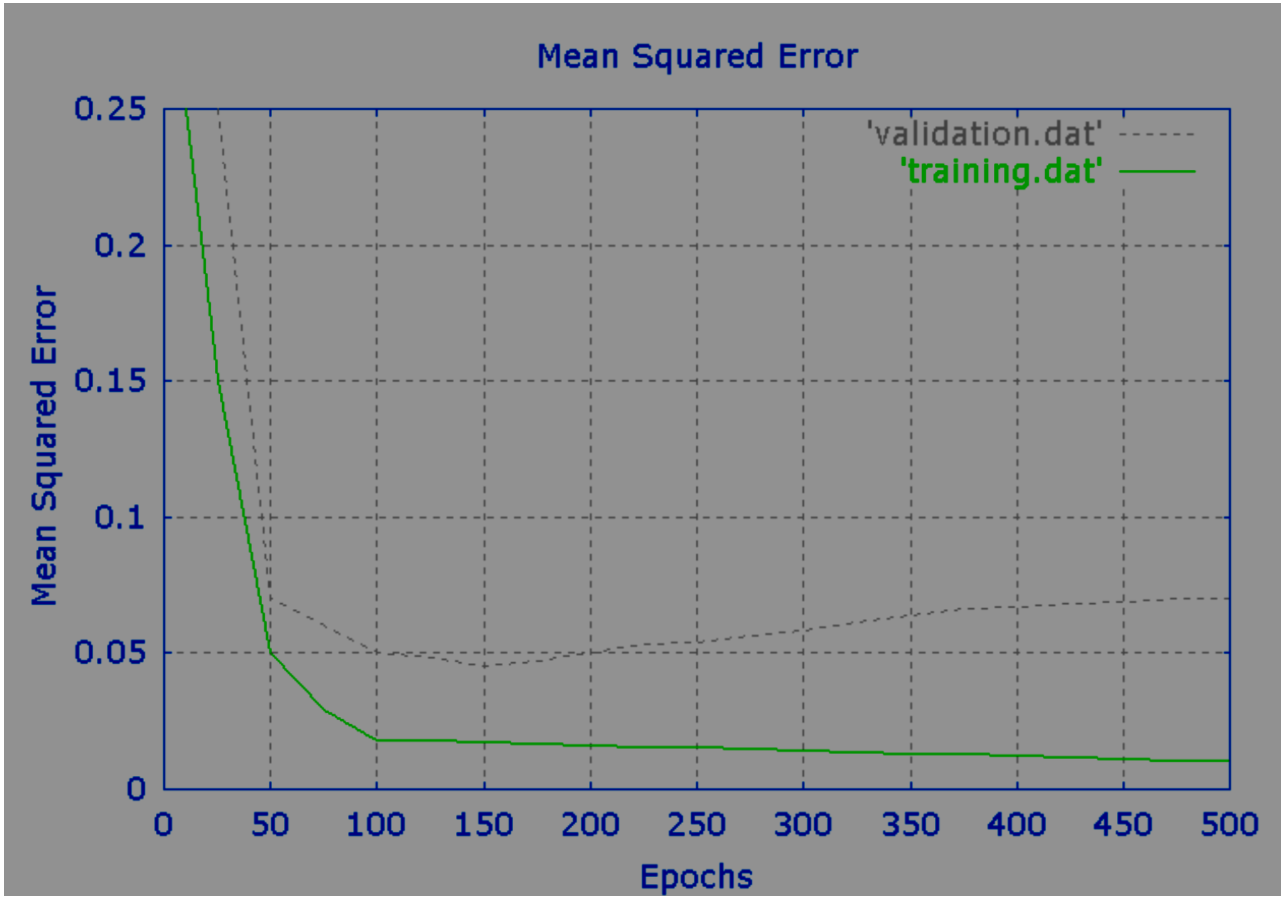


Some problems in ANN applications

- Overfitting
- Data Completeness
- Extrapolation
- The 'ultimate black box'

Not quite

Overfitting Problem



Summary

Why ANN?

- The ability to learn from environment
- The ability to adapt to environment
- Highly non-linear
- Universal approximation (I-O mapping)