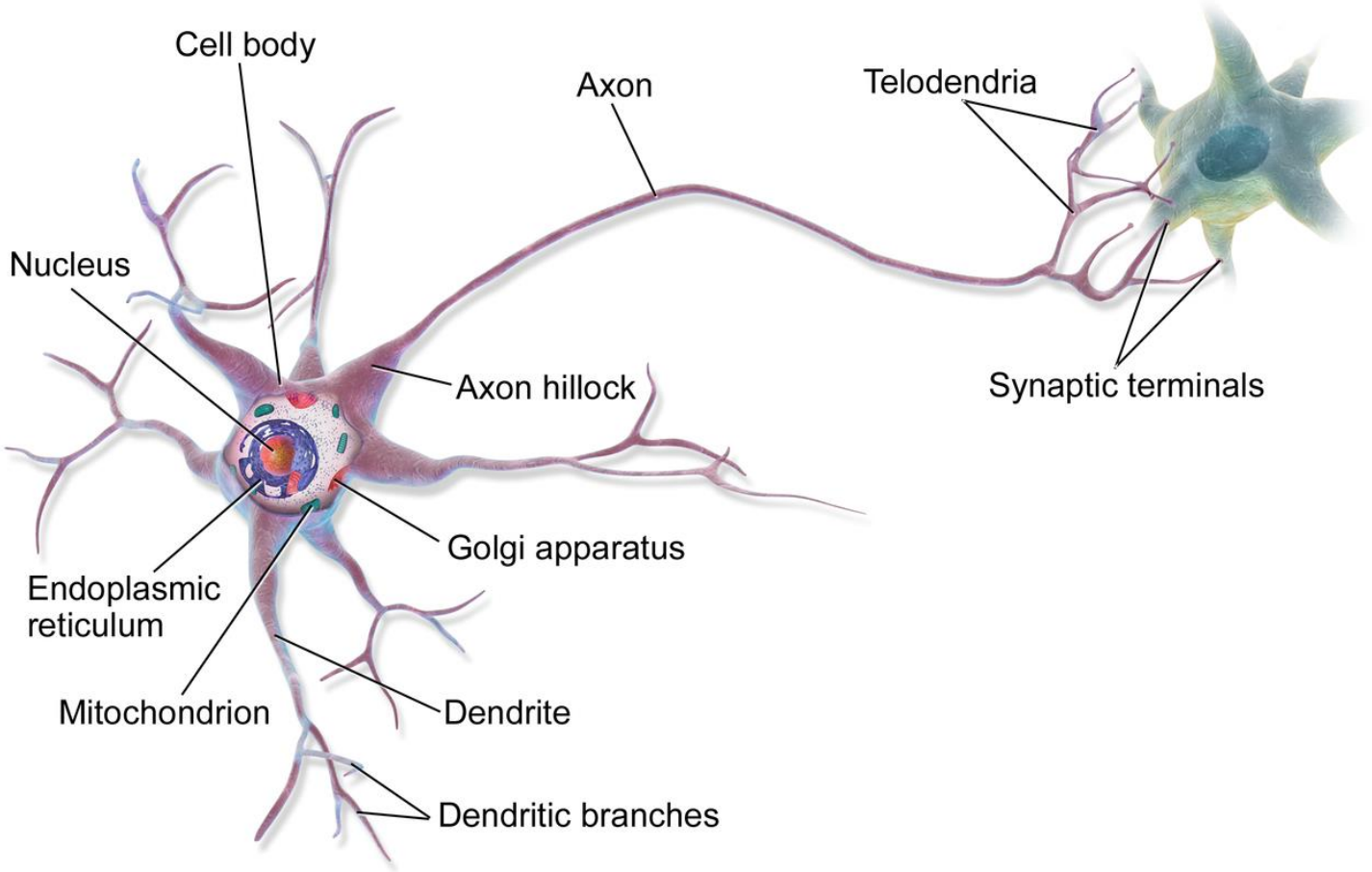
Ural Federal University
named after the first President of Russia B.N.Yeltsin
Institute of radioelectronics and information technologies

Neural Networks

Andrey Sozykin
Andrey.Sozykin@urfu.ru

Ural Federal University
named after the first President of Russia B.N.Yeltsin
Institute of radioelectronics and information technologies

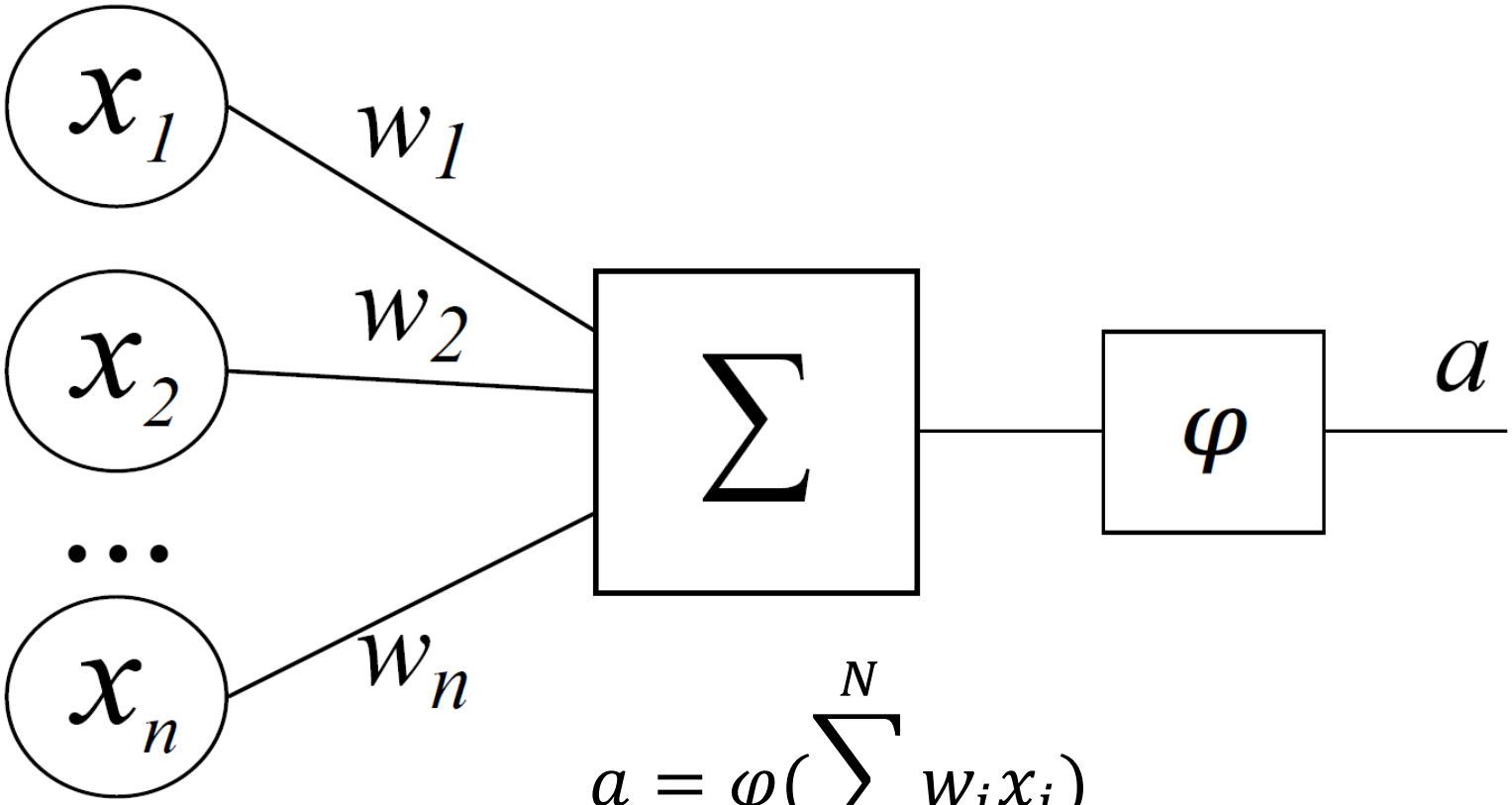
Biological Neuron



Source: https://en.wikipedia.org/wiki/Neuron

Ural Federal University
named after the first President of Russia B.N.Yeltsin
Institute of radioelectronics and information technologies

Artificial Neuron



$$a = \varphi\left(\sum_{i=1}^{N} w_i x_i\right)$$

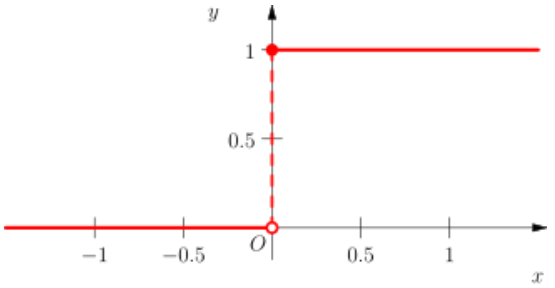
McCulloch and Pitts

Ural Federal University
named after the first President of Russia B.N.Yeltsin
Institute of radioelectronics and information technologies

Activation functions

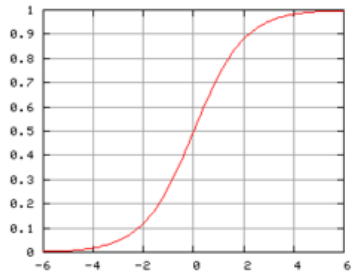
- Heaviside function

$$-\theta(x) = \begin{cases} 0, x < 0 \\ 1, x > 0 \end{cases}$$



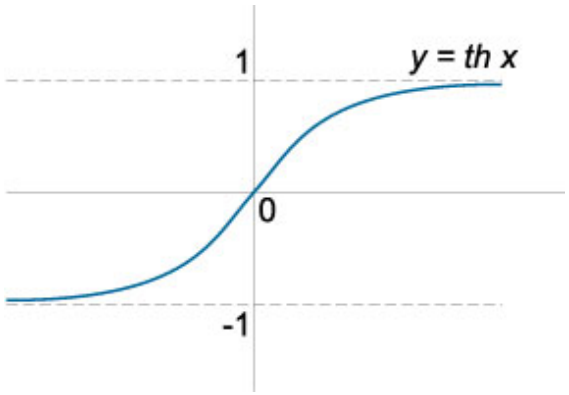
- Sigmoid functions

$$-\sigma(x) = \frac{1}{1+e^{-x}} \text{ (logistics)}$$



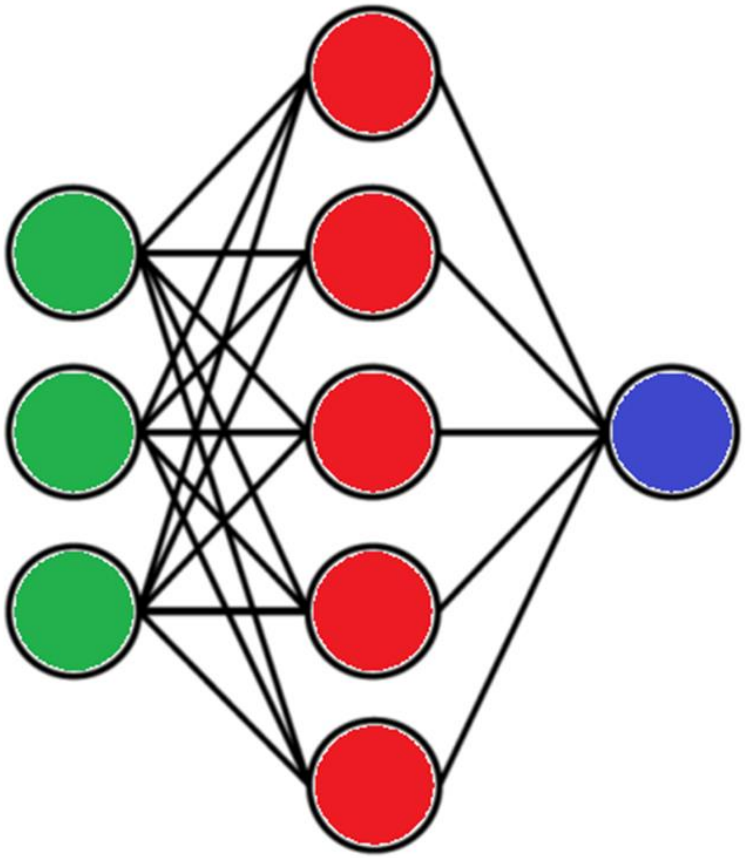
$$-th(x) = \frac{e^{2x}-1}{e^{2x}+1}$$

(hyperbolic tangent)



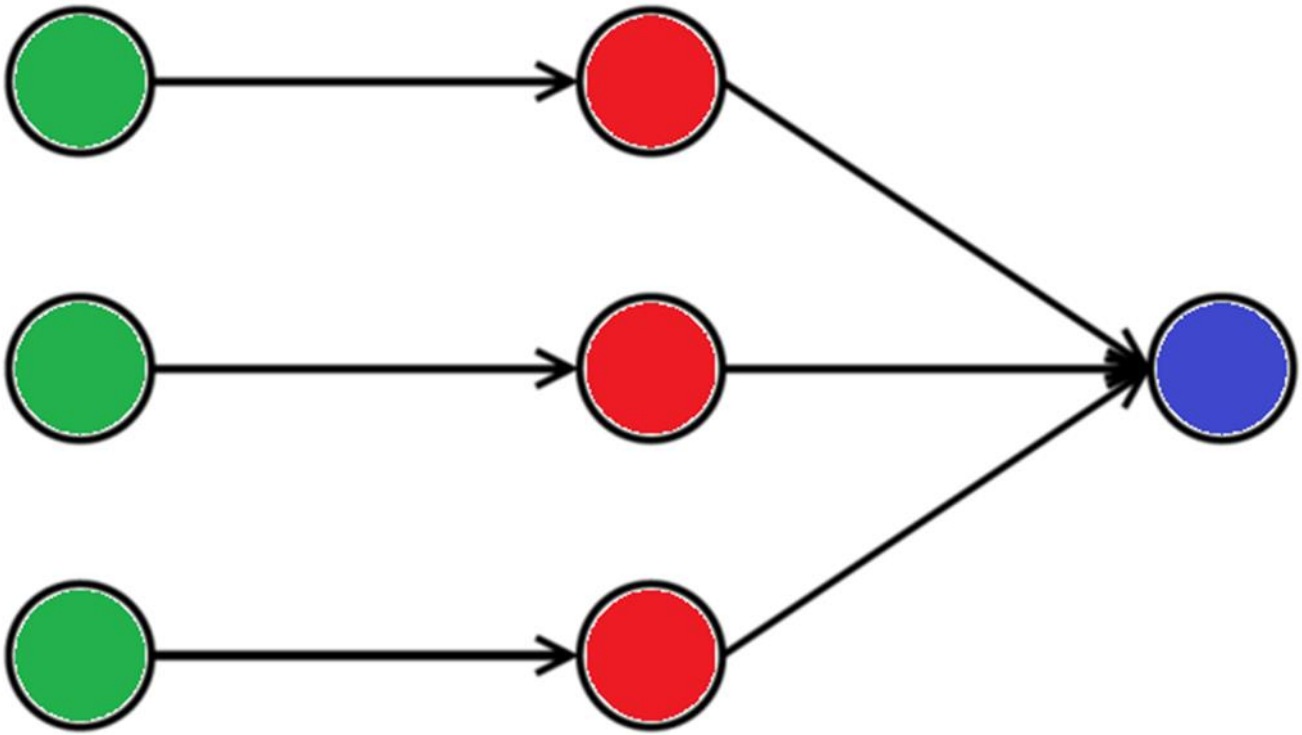
Ural Federal University
named after the first President of Russia B.N.Yeltsin
Institute of radioelectronics and information technologies

Neural Network



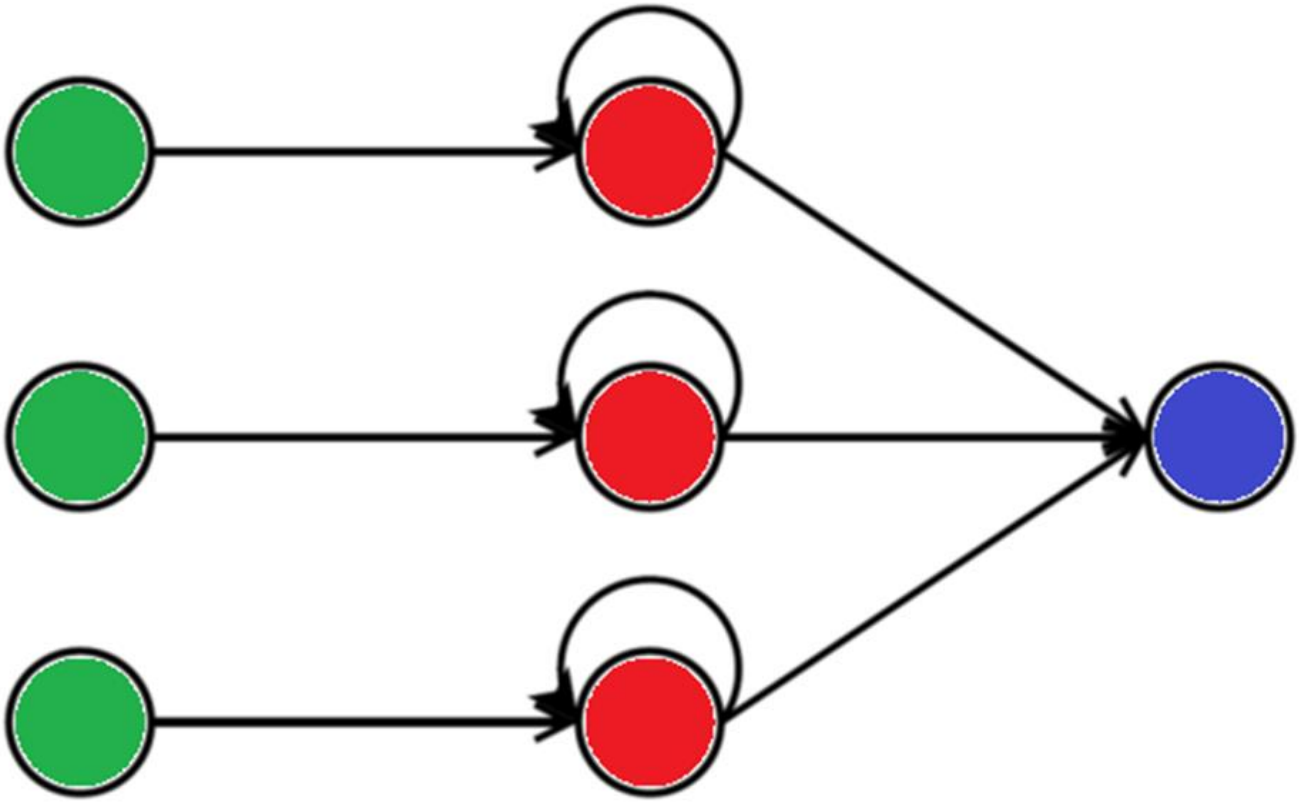
Ural Federal University
named after the first President of Russia B.N.Yeltsin
Institute of radioelectronics and information technologies

Feed Forward Network



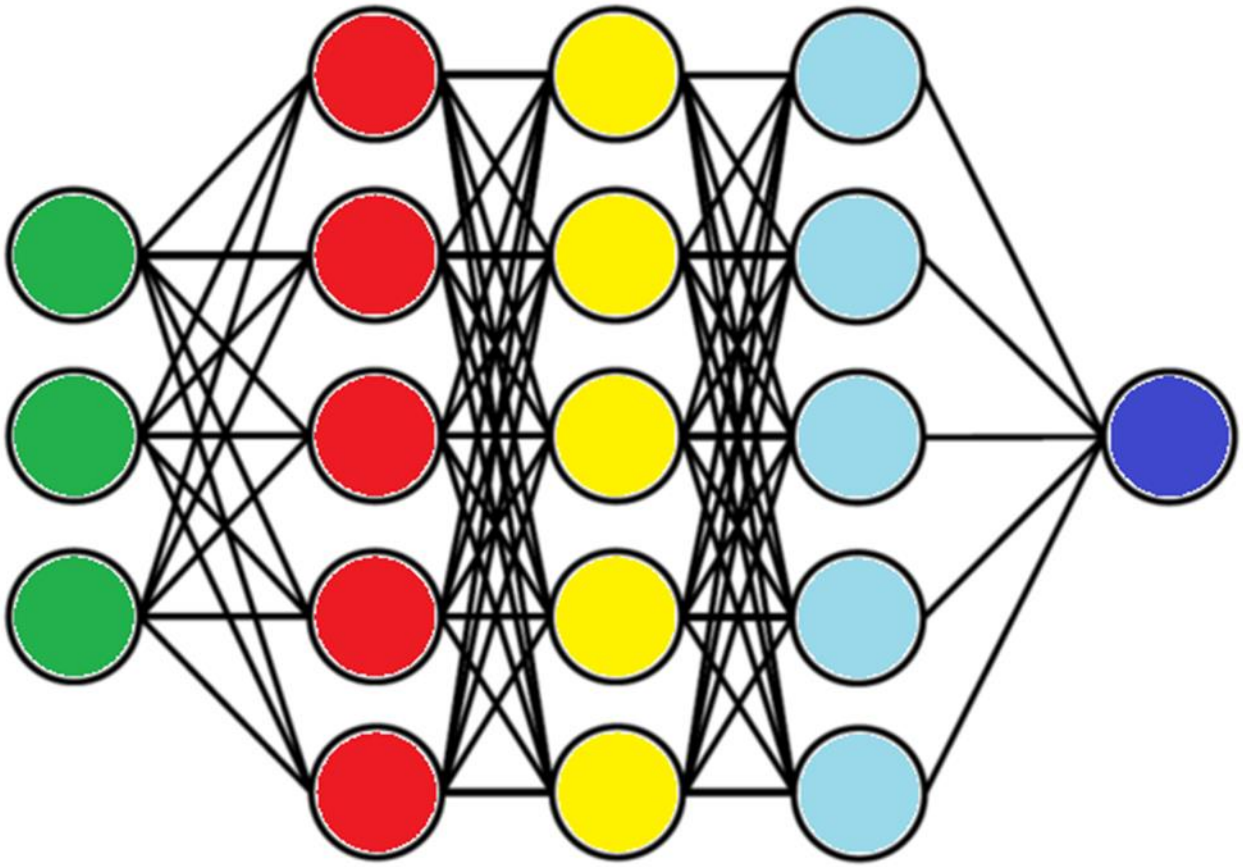
Ural Federal University
named after the first President of Russia B.N.Yeltsin
Institute of radioelectronics and information technologies

Recurrent Neural Network



Ural Federal University
named after the first President of Russia B.N.Yeltsin
Institute of radioelectronics and information technologies

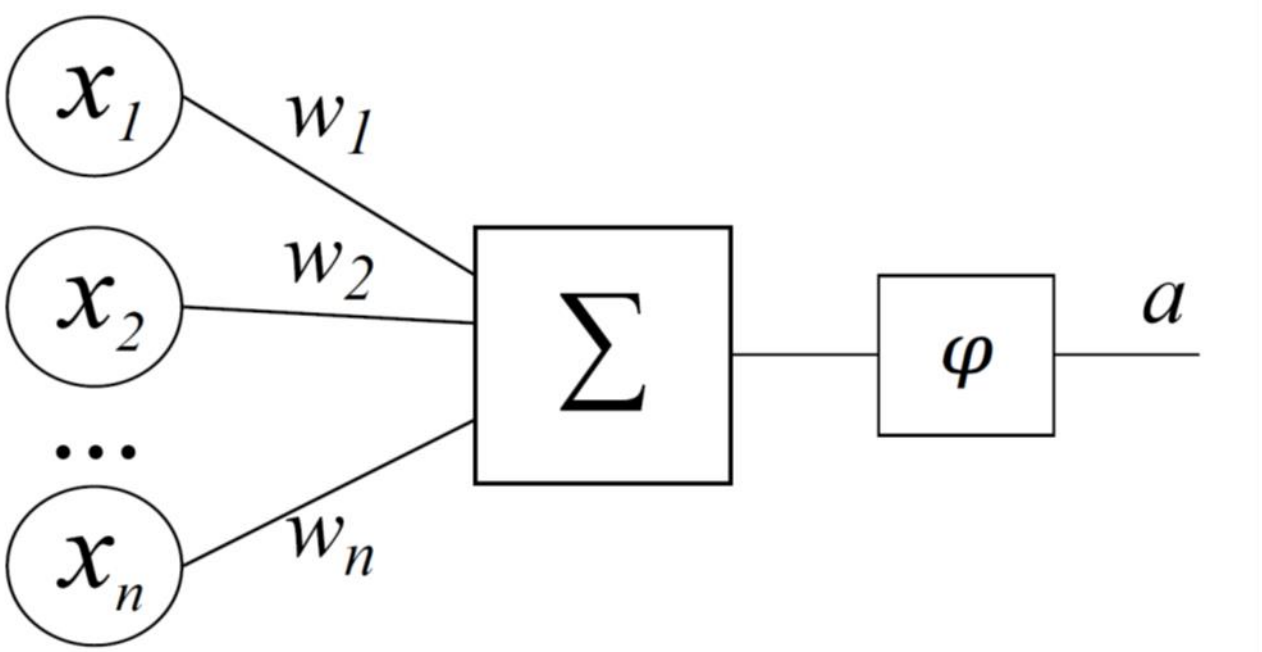
Deep Neural Network



Ural Federal University
named after the first President of Russia B.N.Yeltsin
Institute of radioelectronics and information technologies

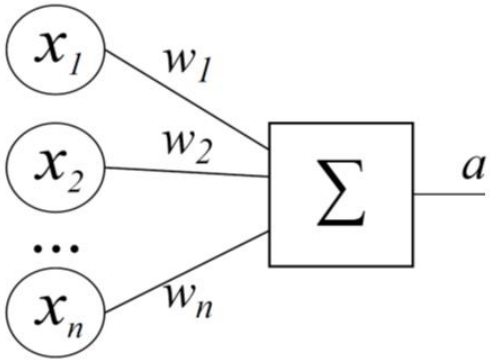
Training Neural Network

Training a model means setting its parameters so that the model best fits the training set.



Ural Federal University
named after the first President of Russia B.N.Yeltsin
Institute of radioelectronics and information technologies

Training Linear Neuron



- Neuron output:

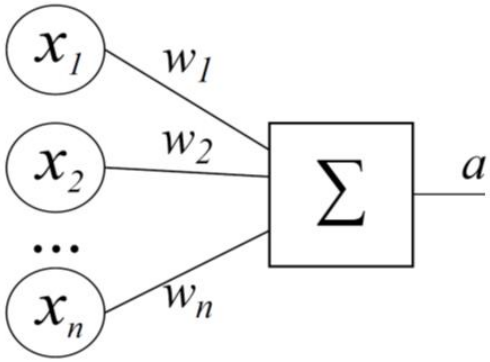
$$a = \sum_{i=1}^{N} w_i x_i$$

- Mean Square Error:

$$\varepsilon = \frac{1}{M} \sum_{j=1}^{M} (a_j - y_j)^2$$

Ural Federal University
named after the first President of Russia B.N.Yeltsin
Institute of radioelectronics and information technologies

Delta rule



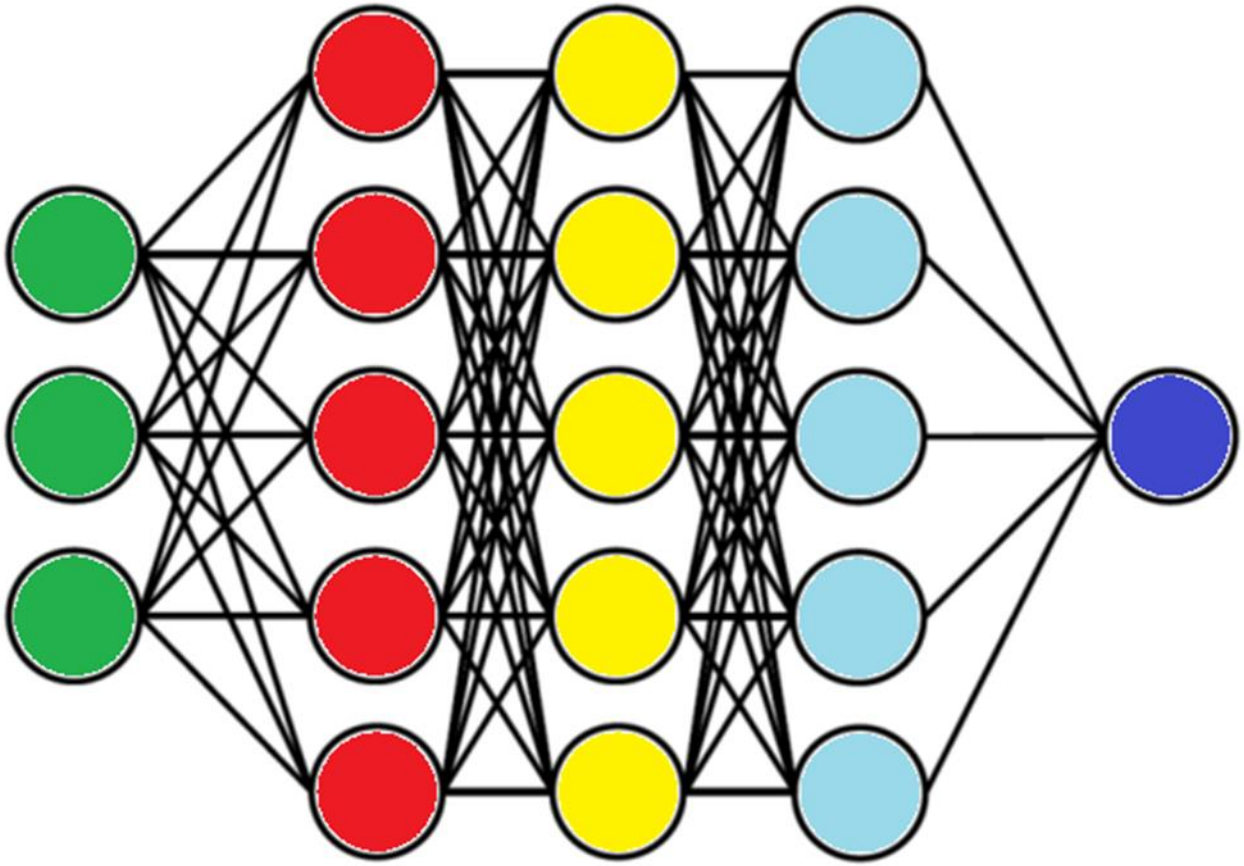
- Changing the weights:

$$w_i = w_i - \eta \frac{2}{M} \sum_{j=1}^{M} x_j^i (a_j - y_j)$$

η – learning rate

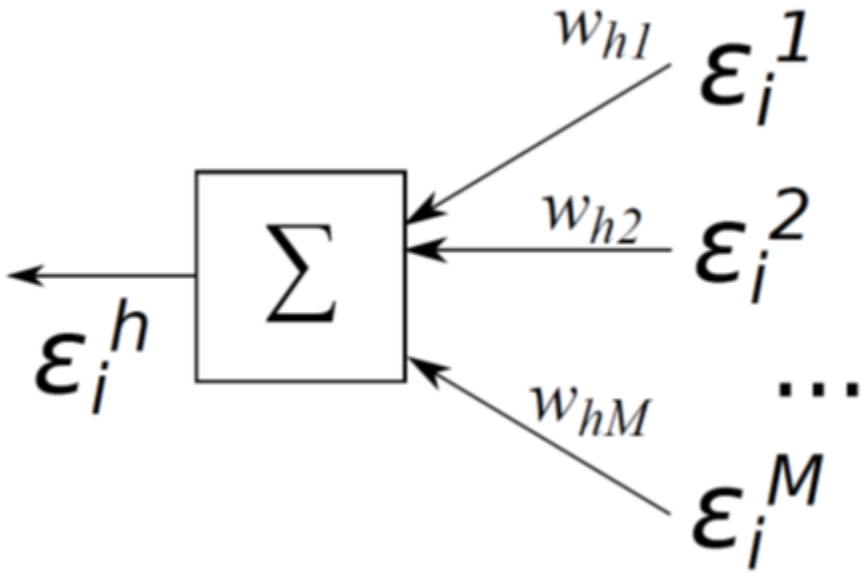
Ural Federal University
named after the first President of Russia B.N.Yeltsin
Institute of radioelectronics and information technologies

Backpropagation Algorithm



Ural Federal University
named after the first President of Russia B.N.Yeltsin
Institute of radioelectronics and information technologies

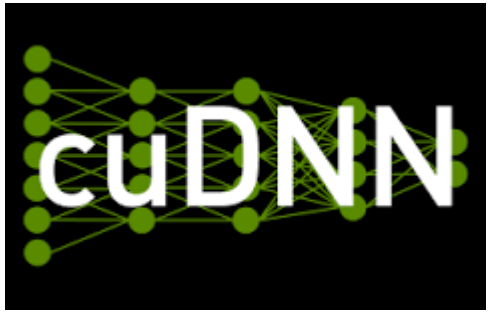
Error Backpropagation



$$\varepsilon_i^h = \sum_{m=1}^{M} \varepsilon_i^m w_{hm}$$

Ural Federal University
named after the first President of Russia B.N.Yeltsin
Institute of radioelectronics and information technologies

Deep Learning Libraries



Ural Federal University
named after the first President of Russia B.N.Yeltsin
Institute of radioelectronics and information technologies

Deep Learning Stack

theano

Microsoft CNTK

Ural Federal University
named after the first President of Russia B.N.Yeltsin
Institute of radioelectronics and information technologies

MNIST Dataset

Mixed National Institute of Standards and Technology database

Back-Propagation Applied to Handwritten Zip Code Recognition / *Y. LeCun, B. Boser, J. S. Denker et al*. 1989

Ural Federal University
named after the first President of Russia B.N.Yeltsin
Institute of radioelectronics and information technologies

Neural Net for MNIST Recognition

- Input layer:
  - 800 neurons
- Output layer:
  - 10 neurons (the number of classes)
- Examples of the neural network architectures for MNIST
  - https://en.wikipedia.org/wiki/MNIST_database

How to prevent overfitting

- Training dataset
  - Dataset for training the neural network
- Test dataset:
  - Dataset for evaluating the performance and generalization of the neural network after the training

Ural Federal University
named after the first President of Russia B.N.Yeltsin
Institute of radioelectronics and information technologies

Parameters and hyperparameters

- Parameters of the neural network (learned during training):
  - Weights of the neuron inputs
- Hyperparameters of the neural network (must be specified by developer):
  - Number of layers in the network
  - Number of neurons in the layers
  - Type of layers (dense, convolutional, recurrent, etc.)
  - Learning rate
  - Number of epoch for training

How to prevent overfitting

- Training dataset
  - Dataset for training the neural network
- Validation dataset
  - Dataset for evaluating the performance and generalization of the neural network **during the training** for tuning **hyperparameters**
- Test dataset:
  - Dataset for evaluating the performance and generalization of the neural network **after the training**

Ural Federal University

named after the first President of Russia B.N.Yeltsin

Institute of radioelectronics and information technologies

Thank you!