4. Machine Learning Overview

4.4 Neural networks

Ricardo Brauner

31 Julho, 2020

Instituto UFC Virtual

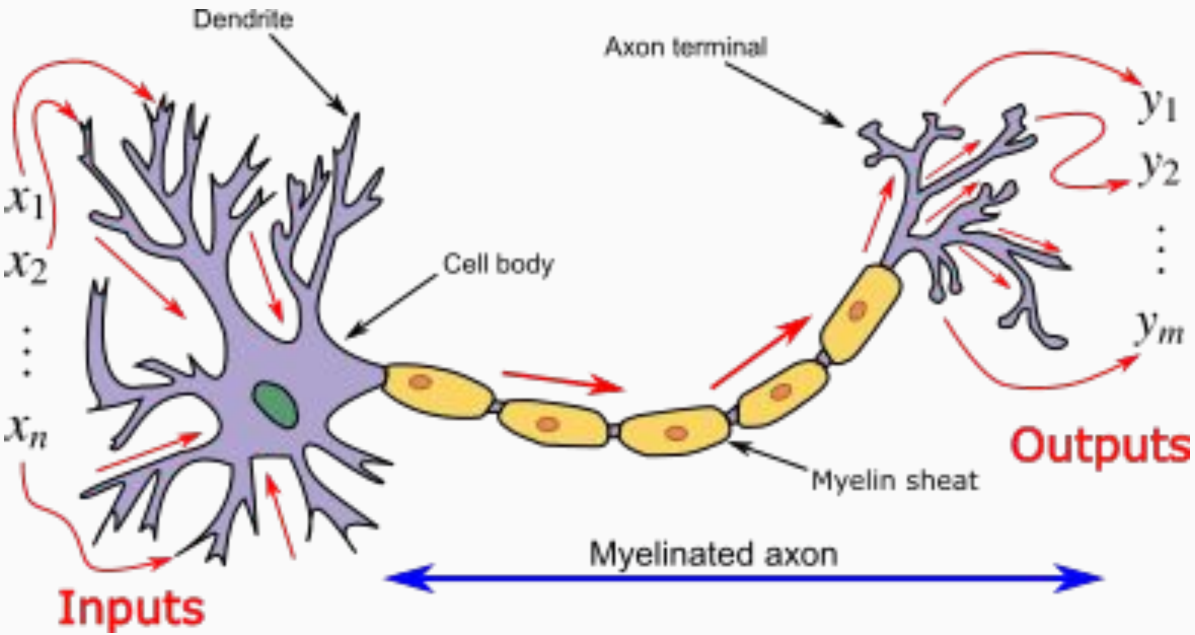
Artificial neural network

- An information processing system
- Designed to simulate human brain
- Composed of artificial neurons
- Simplified human brain

Artificial neural network

- Simulate human intelligence
- Reflects some basic features of human brain:
  - Parallel information processing,
  - Learning,
  - Association,
  - Pattern classification,
  - Memory

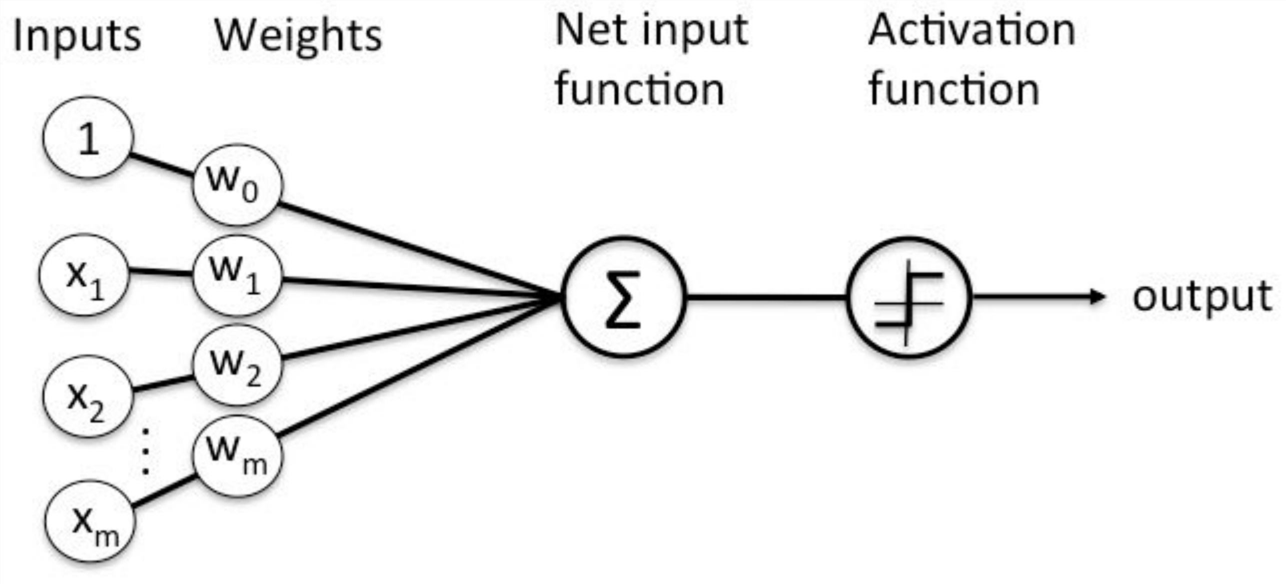
Biological neuron



Artificial Neuron

- Conceptually derived from biological neurons
- Has multiple inputs and produce a single output which can be sent to multiple other neurons
- The weighted sum of all the inputs
- Add a bias term to this sum
- Sum is then passed through a activation function to produce the output

Perceptron



Inputs
Weights
Net input function
Activation function
1
w_0
x_1
w_1
x_2
w_2
x_m
w_m
Σ
output

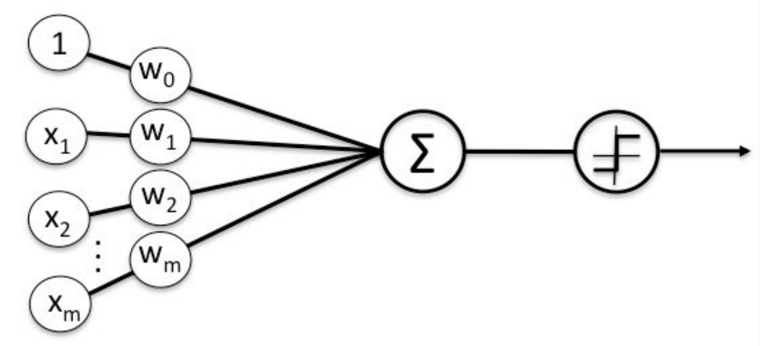
Perceptron

Input vector: $X = [x_0, x_1, \dots, x_n]^T$.

Weight: $W = [\omega_0, \omega_1, \dots, \omega_n]^T$, where ω_0 is the bias.

Activation function: $O = sign(net) = \begin{cases} 1, net > 0, \\ -1, otherwise. \end{cases}$

$$net = \sum_{i=0}^{n} \omega_i x_i = \boldsymbol{W^T X}$$

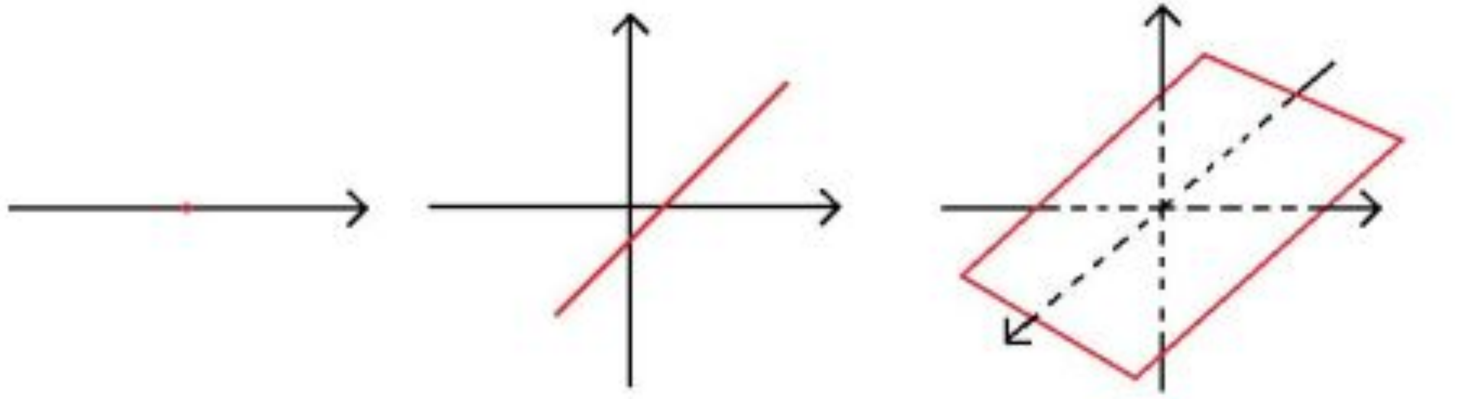


Classifier

- The perceptron is equivalent to a classifier.
- Its input is the high-dimensional vector $\boldsymbol{X}$
- it performs binary classification on input samples in the high-dimensional space.
- If $\boldsymbol{W}^T\mathbf{X} > 0$, the sample is classified into one class.
- Otherwise, the sample is classified into the other class.

Boundary

The boundary $\boldsymbol{W}^T\mathbf{X} = 0$ is a high-dimensional hyperplane.



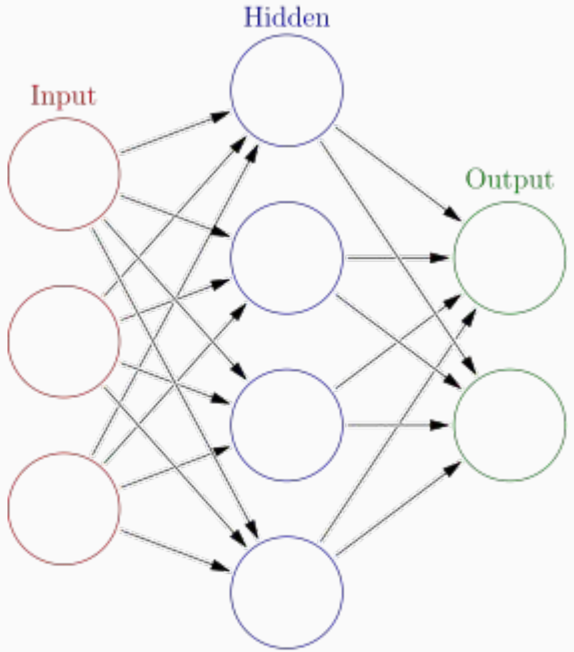
Segmentation point	Segmentation line	Segmentation surface
$Ax + B = 0$	$Ax + By + C = 0$	$Ax + By + Cz + D = 0$

Layers

- Multiple layers
- Especially in deep learning
- Neurons of one layer connect only to neurons of the immediately preceding and immediately following layers

Layers

- The input layer receives external data
- The output layer produces the final result
- In between them are zero or more hidden layers.
- Single layer or single neuron



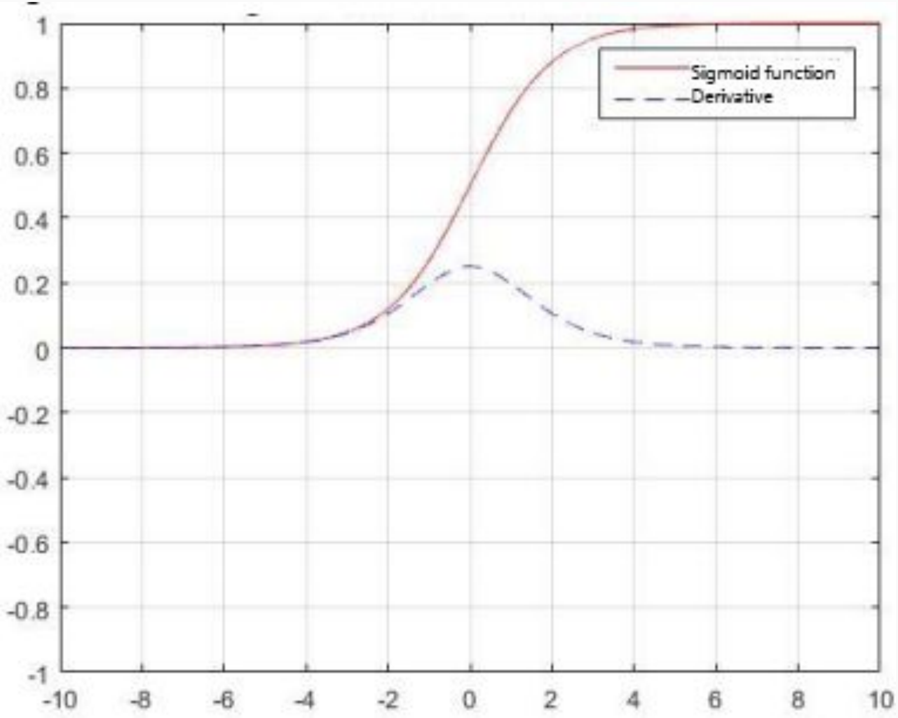
Activation functions

- Non-linear
- If the function is linear, the entire network is equivalent to a single-layer linear model.
- Continuously differentiable
- Enables gradient-based optimization methods.
- Range: When the range of the activation function is finite, gradient-based training methods tend to be more stable.
- When the range is infinite, training is generally more efficient but it is easy to diverge.
- Smaller learning rates are typically necessary.

Activation functions

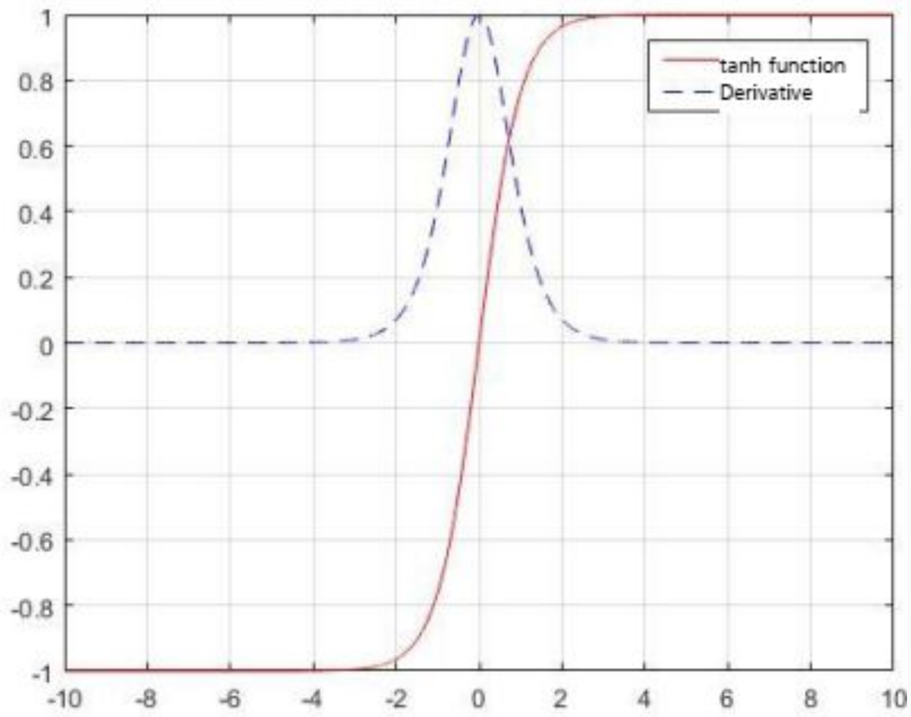
- Monotonicity: The loss function in a single-layer model is convex.
- Smooth: Generalizes better in some cases.
- Approximating identity near the origin: learn efficiently when its weights are initialized with small random values.

Sigmoid function



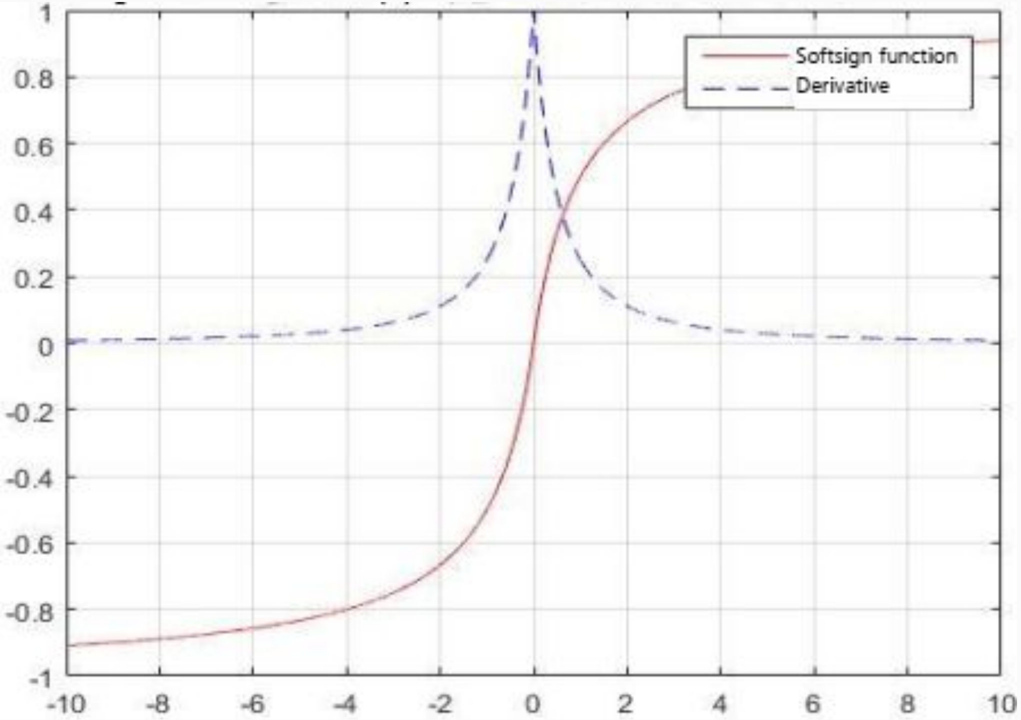
$$f(x) = \frac{1}{1 + e^{-x}}$$

Tanh function



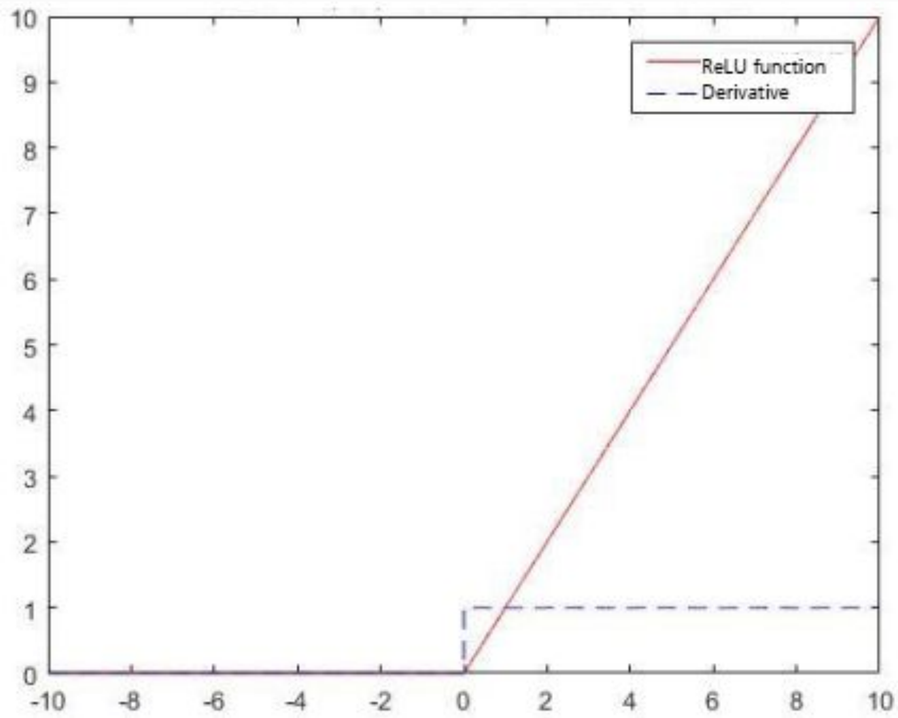
$$\tanh x = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

Softsign function



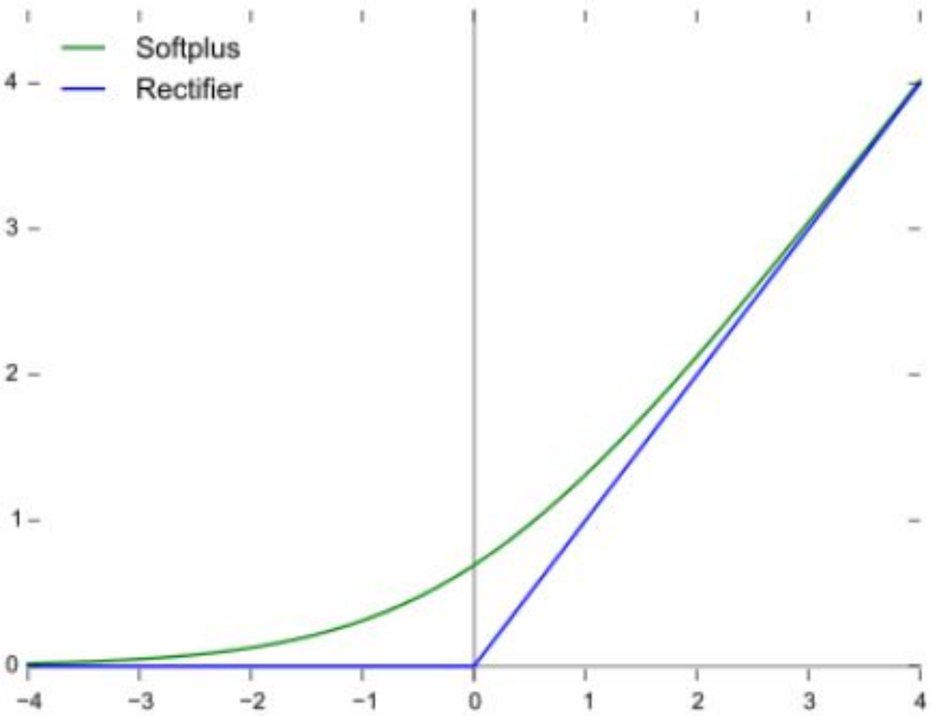
$$f(x) = \frac{x}{|x| + 1}$$

Rectified Linear Unit (ReLU) function



$$y = \begin{cases} x, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

Softplus function



$$f(x) = \ln(e^x + 1)$$

Types of Neural Networks

Single Perceptron

Only represents the linear decision surface

Hyperplane

Multi-layer Fully Connected

- If we connect many perceptrons
- Like a human brain does and
- Non-linear activation function with a function,
- Express a wide range of non-linear surfaces

Feedforward Neural Network

one of the simplest

Widely used

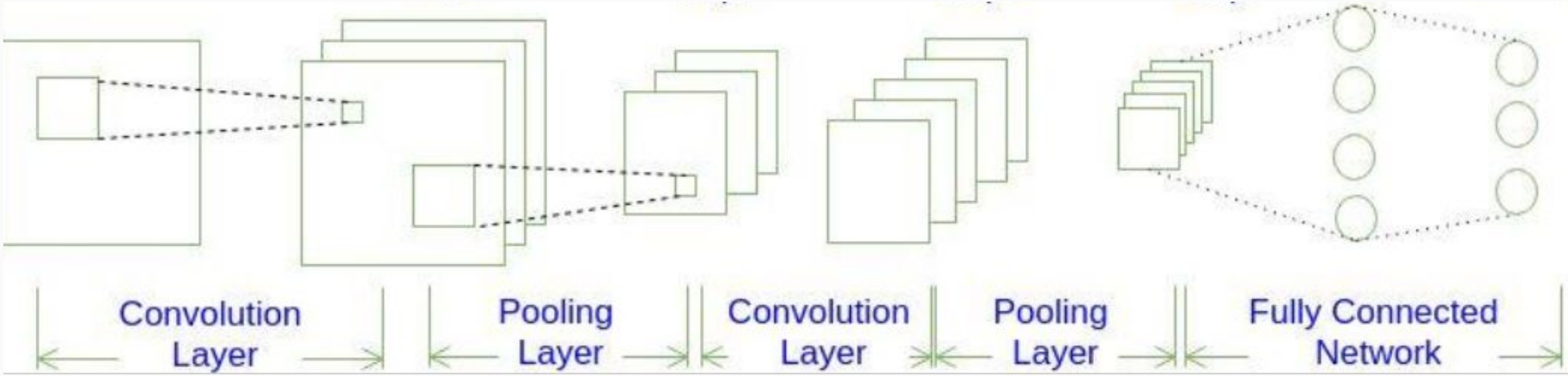
Fastest development

input node does not support computation

Convolutional Neural Network

- Respond to surrounding units
- Image processing
- A convolutional layer: slides a filter on the signal
- A pooling layer: merges info to reduce size
- A fully connected layer: final decision

Convolutional Neural Network



Recurrent Neural Network

- Captures dynamic information in sequential data
- Periodical connections of hidden layer
- Can classify sequential data
- Used in scenarios related to sequences
- Videos, audio and sentences

Recurrent Neural Network

- Neuron output is one of its inputs
- Input can be delayed