AIF183136-03

Week 6 : Neural Network - 1

by Raymond Chandra Putra

INFORMATIKA
UNPAR

Overview

- Intro Neural Network
- Perceptron vs Neuron
- Neuron Type
- Multilayer Perceptron
- Topology
- Feed Forward

What is a neural network ?

The simplest definition of a neural network, more properly referred to as an 'artificial' neural network (ANN), is provided by the inventor of one of the first neurocomputers, Dr. Robert Hecht-Nielsen. He defines a neural network as:

> *"...a computing system made up of a number of simple, highly interconnected processing elements, which process information by their dynamic state response to external inputs.*
>
> In "Neural Network Primer: Part I" by Maureen Caudill, AI Expert, Feb. 1989

ANNs are processing devices (algorithms or actual hardware) that are loosely modeled after the neuronal structure of the mamalian cerebral cortex but on much smaller scales.

http://pages.cs.wisc.edu/~bolo/shipyard/neural/local.html

What is a neural network ? (2)

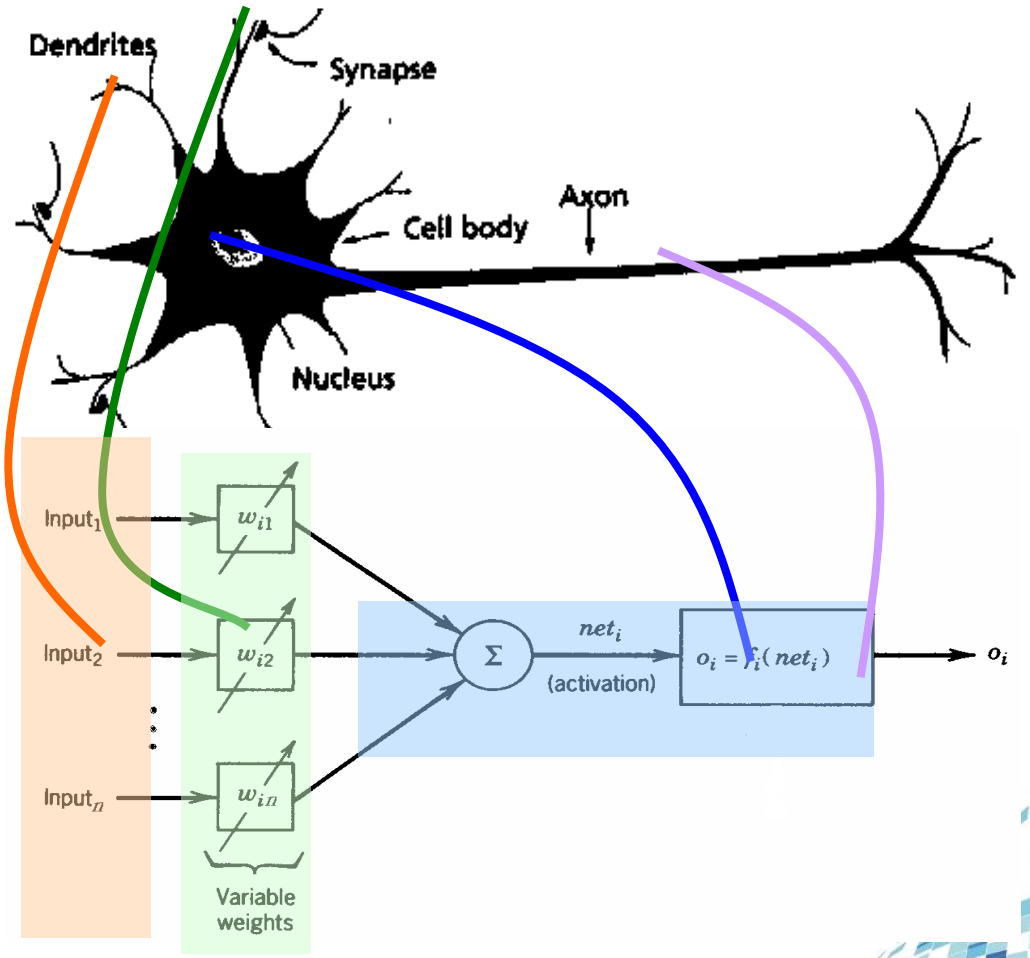
- An interconnected assembly of simple processing elements, units, neurons or nodes, whose functionality is loosely based on the animal neuron.
- The processing ability of the network is stored in the interunit connection strengths, or weights, obtained by a process of adaptation to, or learning from, a set of training patterns.

Where are ANN used?

- Recognizing and matching complicated, vague, or incomplete patterns.
- Data is unreliable
- Problems with noisy data
  - Prediction
  - Classification
  - Data association
  - Pattern Recognition
  - Optimization

Biological vs Artificial

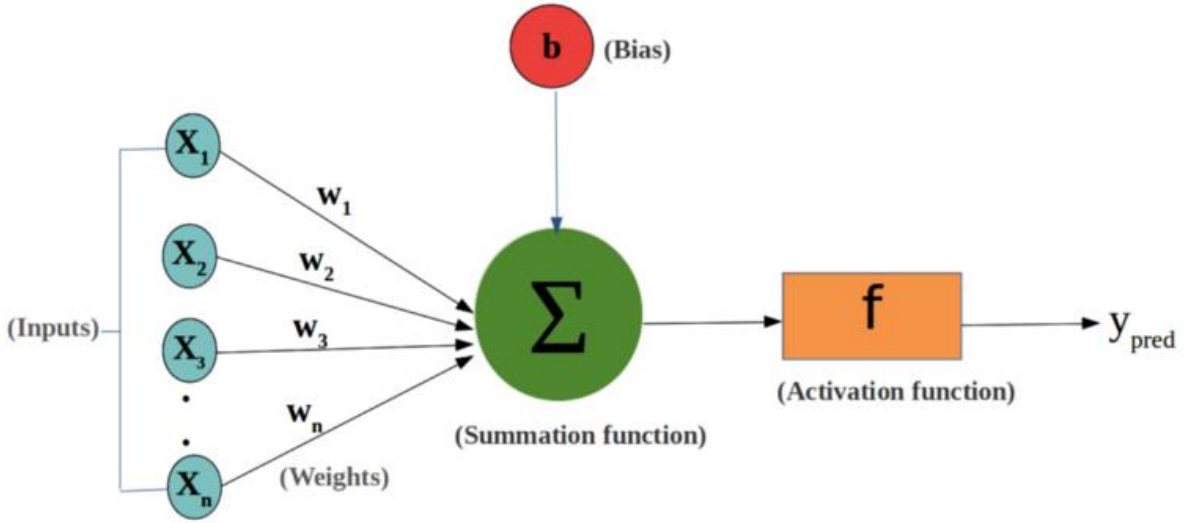
- Biological neuron
- Artificial neuron



INFORMATIKA UNPAR

Artificial Neuron

- Artificial neuron is a modeled after biological neuron (only an approximation).

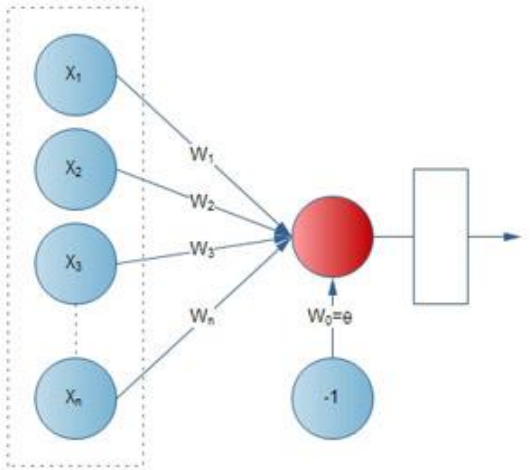


Artificial Neuron Component

- Weights
  - There is a weight associated with each input.
- Bias
  - A threshold weight, $b(w_0)$, associated with a bias value of $x_0(=1)$.
- Summation function
  - Produces the weighted sum of the inputs ($net = w_0x_0 + w_1x_1 + \ldots + w_nx_n$).
- Activation Function
  - An activation function which determines whether the neural unit ‘fires’ or not. This function takes the weighted sum, as its input and outputs a single value

Perceptron

- The perceptron algorithm is also termed the single-layer perceptron.
- The simplest feedforward neural network.
- A network with all inputs connected directly to the outputs.



http://www.codeproject.com/Articles/125346/Single-Layer-Perceptron-as-Linear-Classifier

Artificial Neuron vs Perceptron

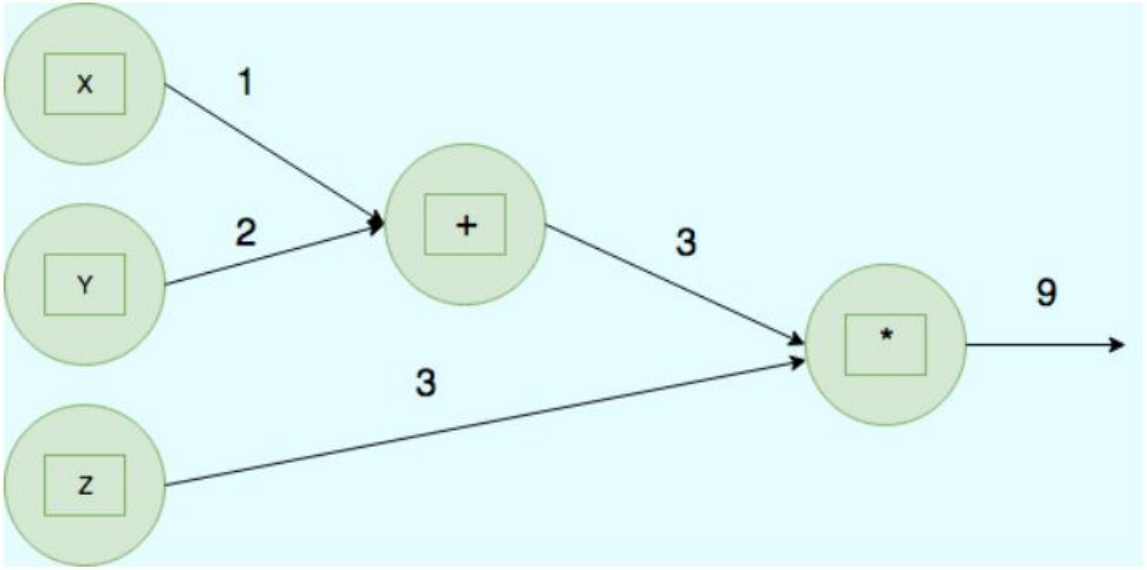
- In the context of neural networks, a perceptron is an artificial neuron using the Heaviside step function as the activation function.

$$f(\mathbf{x}) = \begin{cases} 1 & \text{if } \mathbf{w} \cdot \mathbf{x} + b > 0, \\ 0 & \text{otherwise} \end{cases}$$

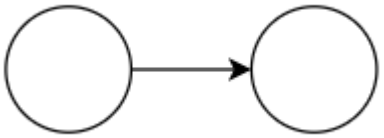
- Means, the results of perceptron are always 0 or 1. But artificial neuron can have other values (also floating number)

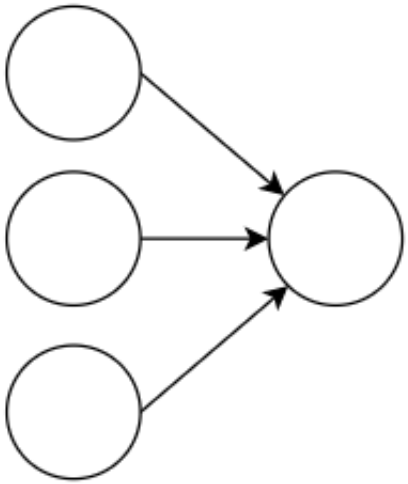
Computational Graph

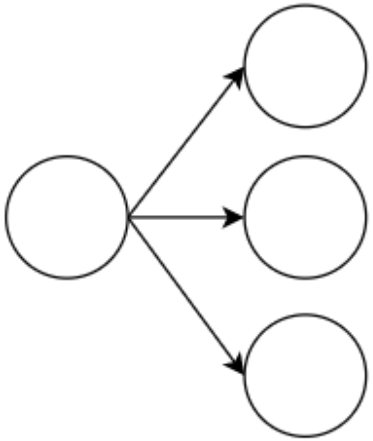
- A computational graph is a way to represent a math function in the language of graph theory.

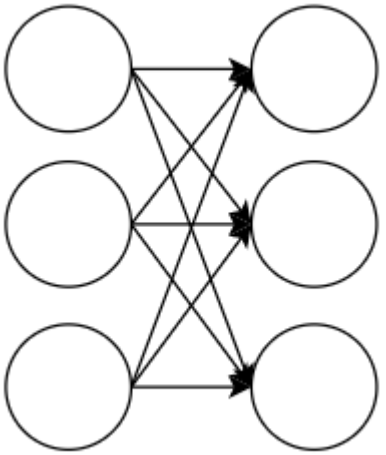


Network Type



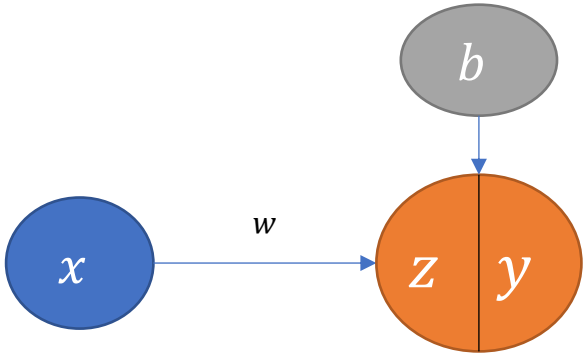






Network Type (1)

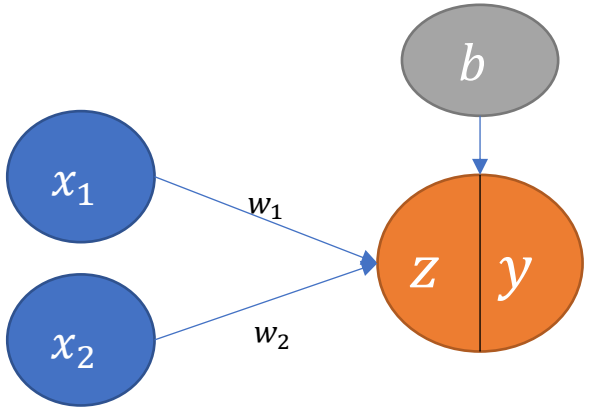
- 1 Input, 1 Output



$$z = wx + b$$
$$y = \sigma(z)$$
$$y = \sigma(wx + b)$$

Network Type (2)

- Multiple Input, 1 Output



$$z = (w_1 x_1 + w_2 x_2 + \ \ldots) + b$$

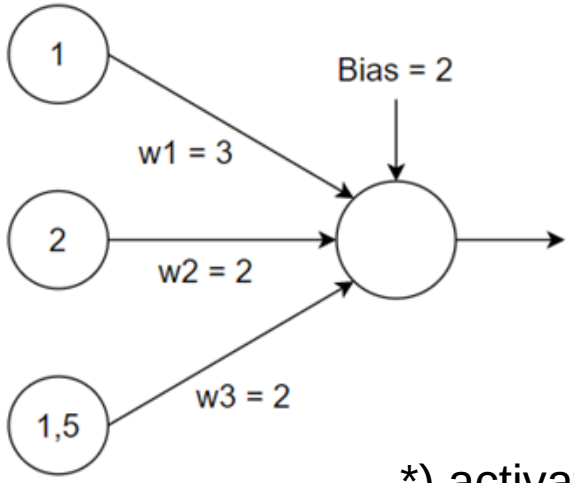
$$y = \ \sigma(z)$$

$$y = \sigma((w_1 x_1 + w_2 x_2 + \ \ldots) + b)$$

$$y = \sigma\left(\sum_{i=1}^{n} w_i \, x_i + b\right)$$

Example:

- The input for the next layer is the weighted sum of all previous layers + bias with activation function.
- Formula : $y = \sigma((x_1 w_1 + x_2 w_2 + \ ... + x_n w_n) + b)$



$$y = (w_1 x_1 + w_2 x_2 + w_3 x_3) + b$$

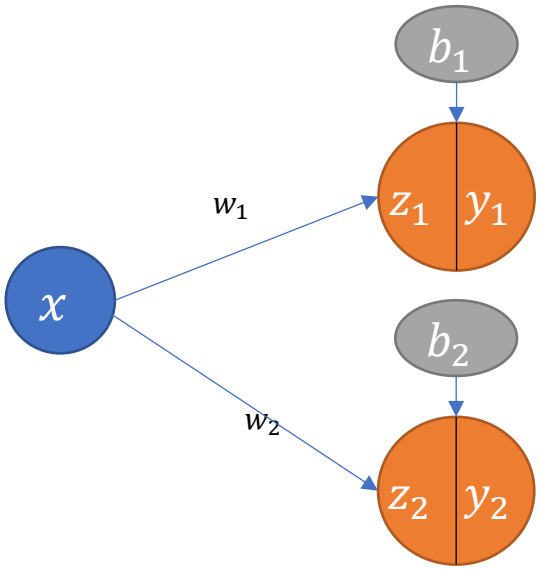
$$y = (3 \times 1 + 2 \times 2 + 2 \times 1.5) + 2$$

$$y = 12$$

*) activation function is ignored

Network Type (3)

- 1 Input, Multiple Output



$$y_1 = \sigma(w_1 x + b_1)$$
$$y_2 = \sigma(w_2 x + b_2)$$

$$\hat{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \sigma(w_1 x + b_1) \\ \sigma(w_2 x + b_2) \end{bmatrix}$$

$$\hat{y} = \sigma(\vec{w}\, x + \vec{b}\,)$$

$$\vec{w} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \qquad \vec{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

Vectorized Function

$$\hat{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \sigma(w_1 x + b_1) \\ \sigma(w_2 x + b_2) \end{bmatrix}$$

$$\hat{y} = \sigma(\begin{bmatrix} w_1 x + b_1 \\ w_2 x + b_2 \end{bmatrix})$$

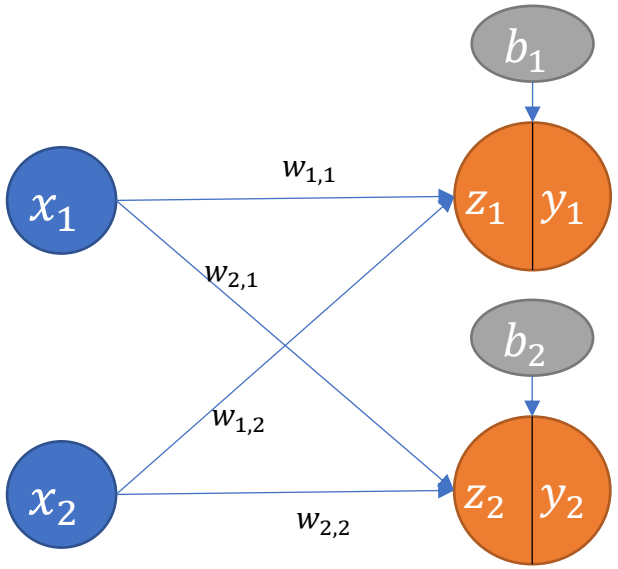
$$\hat{y} = \sigma(\begin{bmatrix} w_1 \\ w_2 \end{bmatrix} x + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix})$$

$$\hat{y} = \sigma(\vec{w}\, x + \vec{b}\,)$$

$$f(x, y) = \sigma(\begin{bmatrix} x \\ y \end{bmatrix}) = \begin{bmatrix} \sigma(x) \\ \sigma(y) \end{bmatrix}$$

Network Type (4)

- Multiple Input, Multiple Output



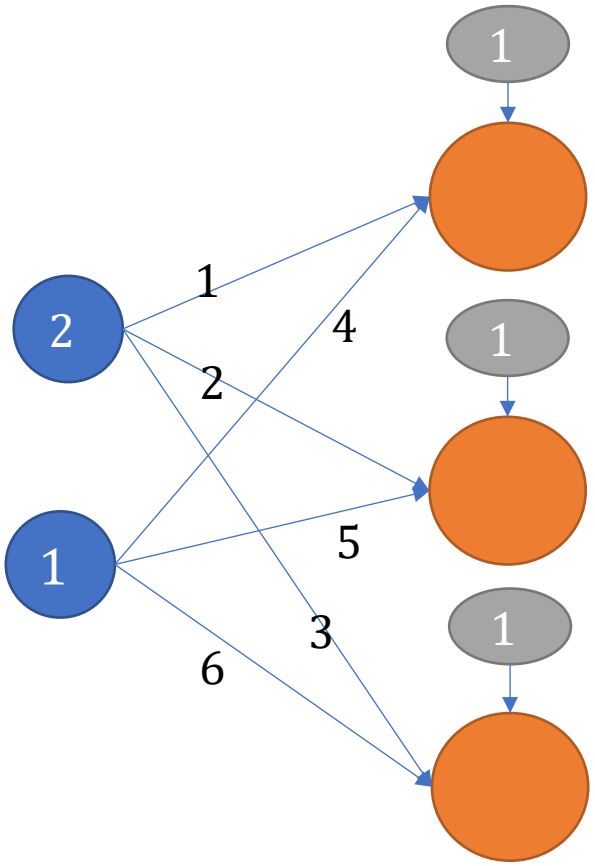
$$\hat{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \sigma(\sum_{i=1}^{n} w_{1,i}x_i + b_1) \\ \sigma(\sum_{i=1}^{n} w_{2,i}x_i + b_2) \end{bmatrix}$$

$$\hat{y} = \sigma(W\vec{x} + \vec{b})$$

$$W = \begin{bmatrix} w_{1,1} & w_{1,2} \\ w_{2,1} & w_{2,2} \end{bmatrix} \quad \vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \vec{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

INFORMATIKA UNPAR

Practice :



- Write on Matrix & vector form.
- Define formula $\hat{y}$.
- Result of $\hat{y}$ = ?

baru ngerjain satu soal matematika

me:

Batch Processing

$$\hat{y} = \sigma(W\vec{x} + \vec{b})$$

- This formula only applied under the assumption data are inputted one at a time.
- Batch processing is the processing of large volume of data all at once.

$$X = \begin{bmatrix} x_1 & x_2 \\ x_1 & x_2 \\ x_1 & x_2 \end{bmatrix}$$

Batch 1
Batch 2
Batch 3

- Formula for batch processing :

$$\hat{Y} = \sigma(WX^T + \vec{b})$$

$$\hat{Y}^T = \sigma(XW^T + \vec{b})$$

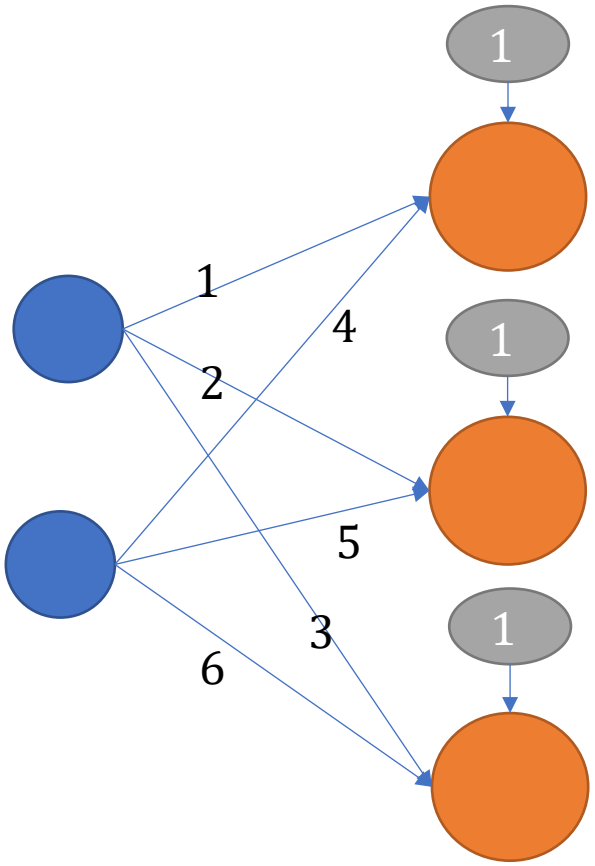
$$Y = WX + b$$
$$Y^T = X^T W^T + b^T$$

INFORMATIKA UNPAR

Math vs Programming

- Code cannot differentiate column vector and row vector.
- Depends on the axis representation, we need to transpose the matrix.
  - Different dataset have different representation (that's why data preparation is important)

Practice :

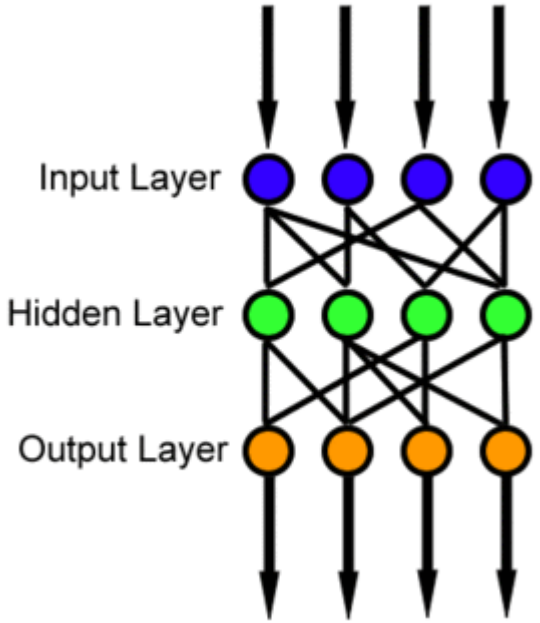


Batch	x1	x2
1	2	1
2	3	5
3	6	6

- Result of $\hat{Y} = ?$

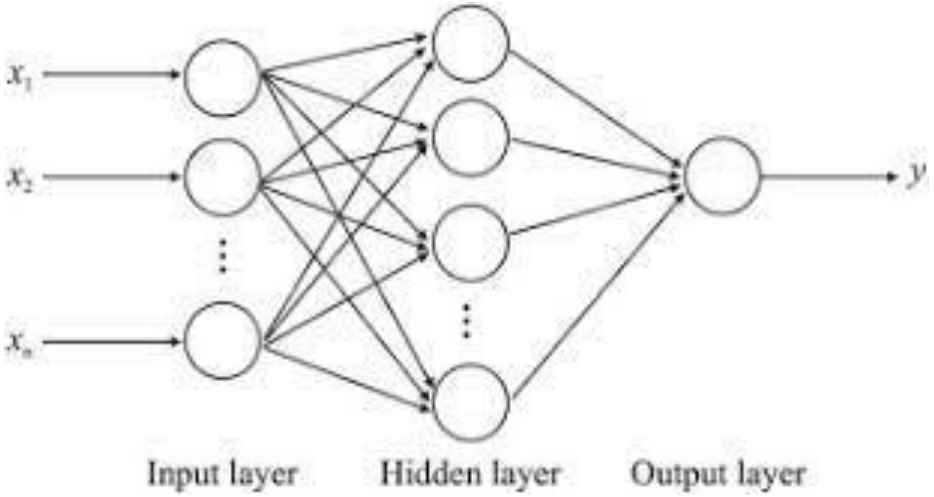
Feed Forward

- A feedforward neural network is an artificial neural network wherein connections between the nodes do not form a cycle.



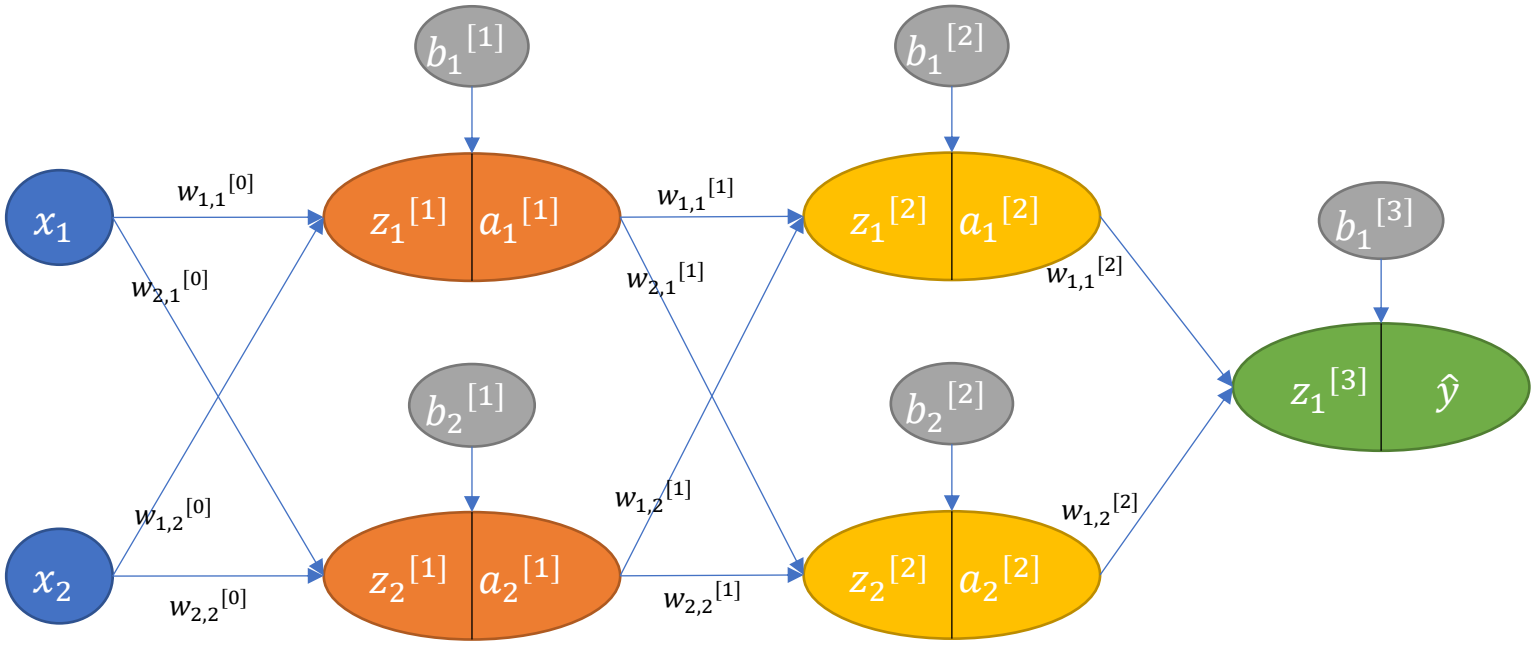
Multilayer Perceptron

- Neural Network with multiple layers.



- Introducing hidden layer (layer in between input layers and output layers).

Multilayer Perceptron

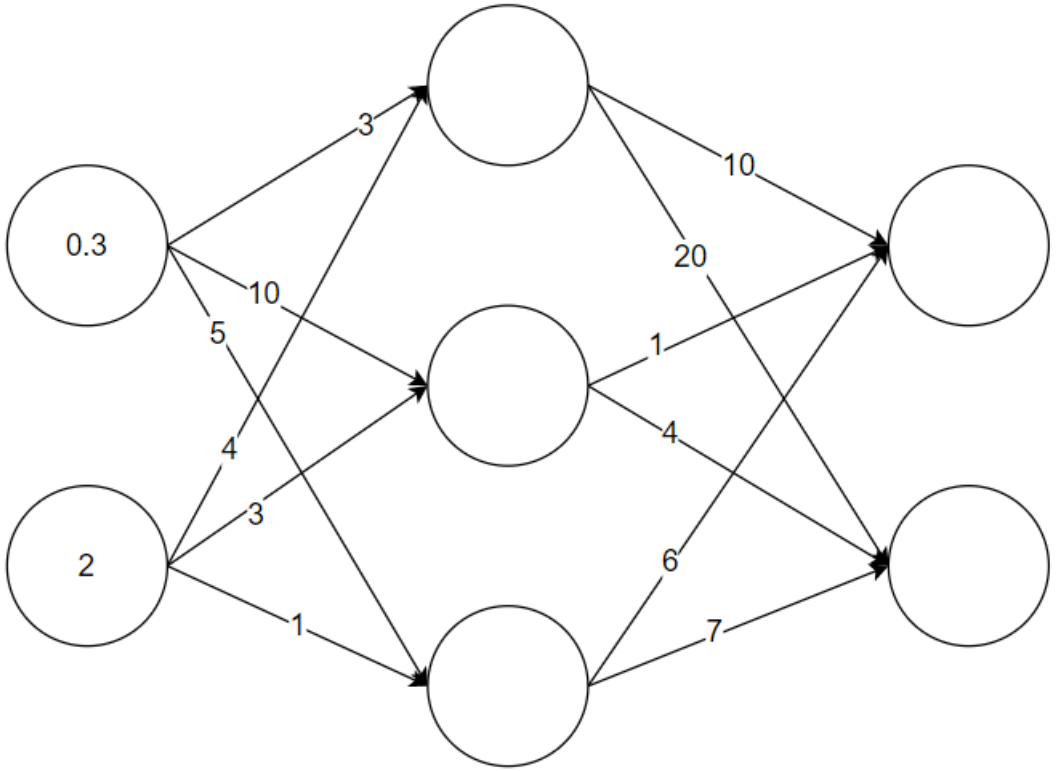


Input Layer Hidden Layer 1 Hidden Layer 2 Output Layer

INFORMATIKA UNPAR

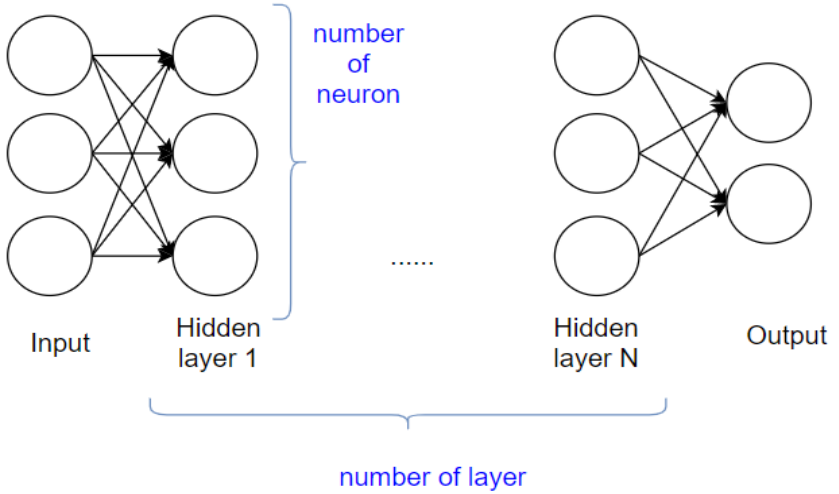
Practice :

- Assume : bias = 0, activation function is ignored.



Neural Network Topology

- Define number of layer and number of neuron in each layers.



- Example :
- Network [3,3,5,2]
  - 4 Layers (1 Input, 2 Hidden, 1 Output)
  - Input Layer : 3 neuron
  - Hidden Layer 1 : 3 neuron
  - Hidden Layer 2 : 5 neuron
  - Output Layer : 2 neuron

Thanks

INFORMATIKA
UNPAR

INFORMATIKA
UNPAR

informatika.unpar.ac.id
informatika@unpar.ac.id
if.unpar
if.unpar