Intelligent Control

Artificial Neural Network (4c)

by
Dr. Nor Maniha Abdul Ghani
FKEE
normaniha@ump.edu.my

Chapter Description

At the end of this topic , student should be able to:-

- Understand the concept of Neural Network and analyse given network using backpropagation algorithm.

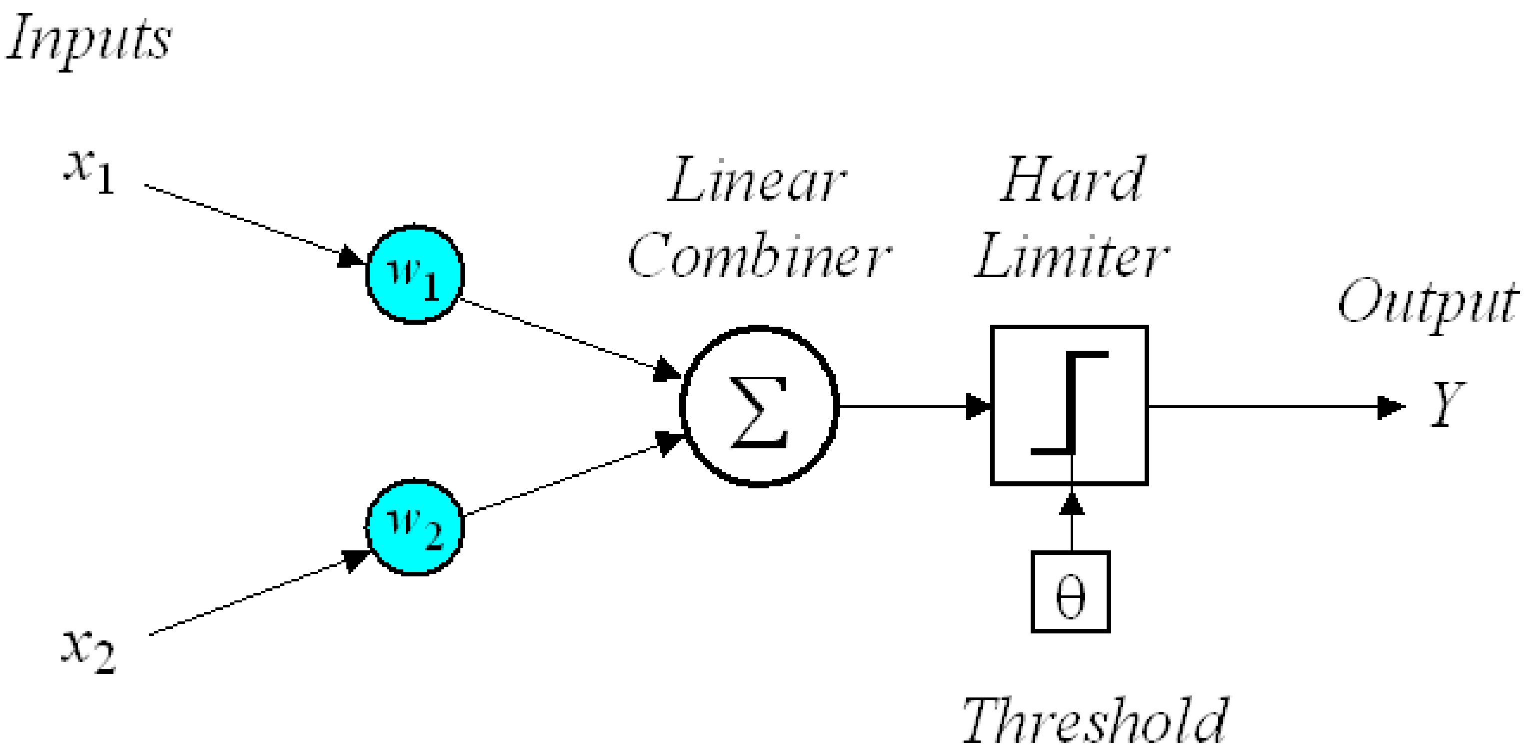
Contents

- **4.5 Simple ANN**
- **4.6 Multilayer Neural Networks & Backpropagation Algorithm**

Simple ANN

4.5

Simple Perceptron



http://slideplayer.com

Learning: Classification

- Learning is done by adjusting the actual output Y to meet the desired output Y_d.
- Usually, the initial weight is adjust between -0.5 to 0.5. At iteration *k* of the training example, we have the error *e* as

$$e(k) = Y_d(k) - Y(k)$$

- If the error is positive, the weight must be decrease and otherwise must be increase.
- Perceptron learning rule also can be obtained where

$$w_i(k+1) = w_i(k) + \alpha \times x_i(k) \times e(k)$$

α is the learning rate and 0< α<1.

Training algorithm

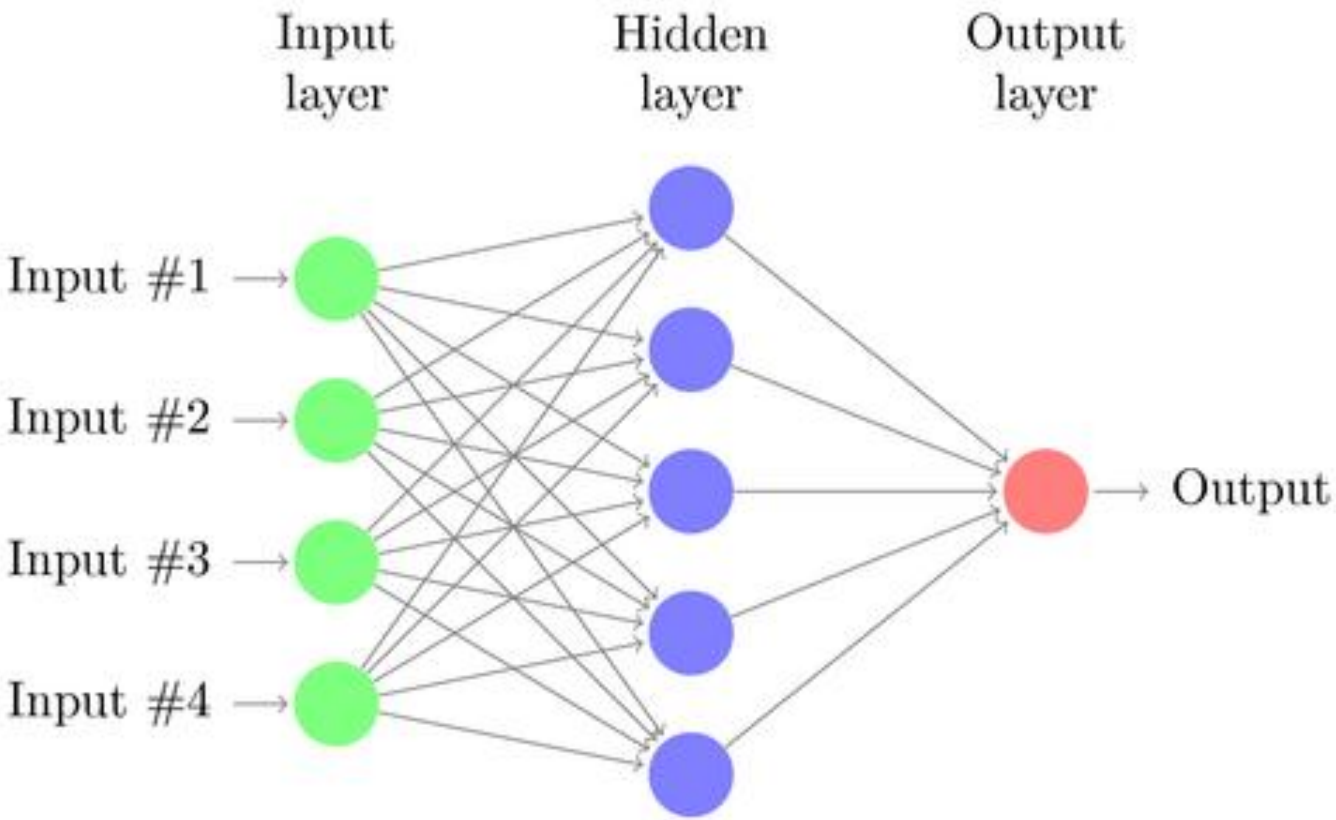
- Step 1: Initialization
  - Set initial weight w_i between [-0.5,0.5]and θ.
- Step 2: Activation
  - Perceptron activation at iteration 1 for each input and a specific Y_d. e.g for a step activation function we have
  - $Y(p) = step[\sum_{i=1}^{n} x_i\,(p) w_i(p) - \theta]$
- Step 3: Weight training
  - Perceptron weight is updated by
  - $w_i(p+1) = w_i(p) + \Delta w_i(p)$

    where $\Delta w_i(p) = \alpha + x_i(p) + e(p)$
- Step 4: Iteration
  - Next iteration at time $k+1$ and go to step 2 again.

Multilayer Neural Networks & Backpropagation Algorithm

4.6

Multilayer neural networks



http://www.texample.net

Back Propagation NN

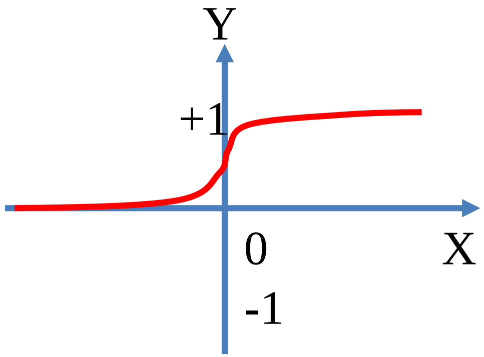
Most popular training algorithm.

Each neuron connect to each other.

Sigmoid function.

Input layer propagates to another layer.

Weight will update according to error value.



$$Y = \frac{1}{1 + e^{-X}}$$

Step of Back Propagation Method

1. Weight is set up.
2. Input pattern is applied.
3. Output vs Target is calculated.
4. Compute Error and update weight.

When does the training process stop?

- Sum squared error for Output, Y is less than Input.
- The smaller sum squared error, the better system performance.
- As indicator of the system

More about back-propagation

Backpropagation	
	Different solution for different weight and threshold.
	Sign activation function for viewing decision boundaries.
	Not suitable for biological neurons.
	Expensive computation.
	Training is slow.

XOR PROBLEM

- In this example we use the BP algorithm to solve a 2-bit XOR problem.
- The training patterns of this ANN is the XOR example as given in the table.
- For simplicity, the ANN model has only 4 neurons (2 inputs, 1 hidden and 1 output) and has no bias weights.
- The input neurons have linear functions and the hidden and output neurons have sigmoid functions.
- The weights are initialised randomly.
- We train the ANN by providing the patterns #1 to #4 through an iteration process until the error is minimized.

	Inputs		Target
	O_{i1}	O_{i2}	t_k
Pattern#1 →	0	0	0
Pattern#2 →	0	1	1
Pattern#3 →	1	0	1
Pattern#4 →	1	1	0

32

Example of Examination question

Figure 1 below illustrates a multilayer Neural Network that has the input patterns of (0 1 1).

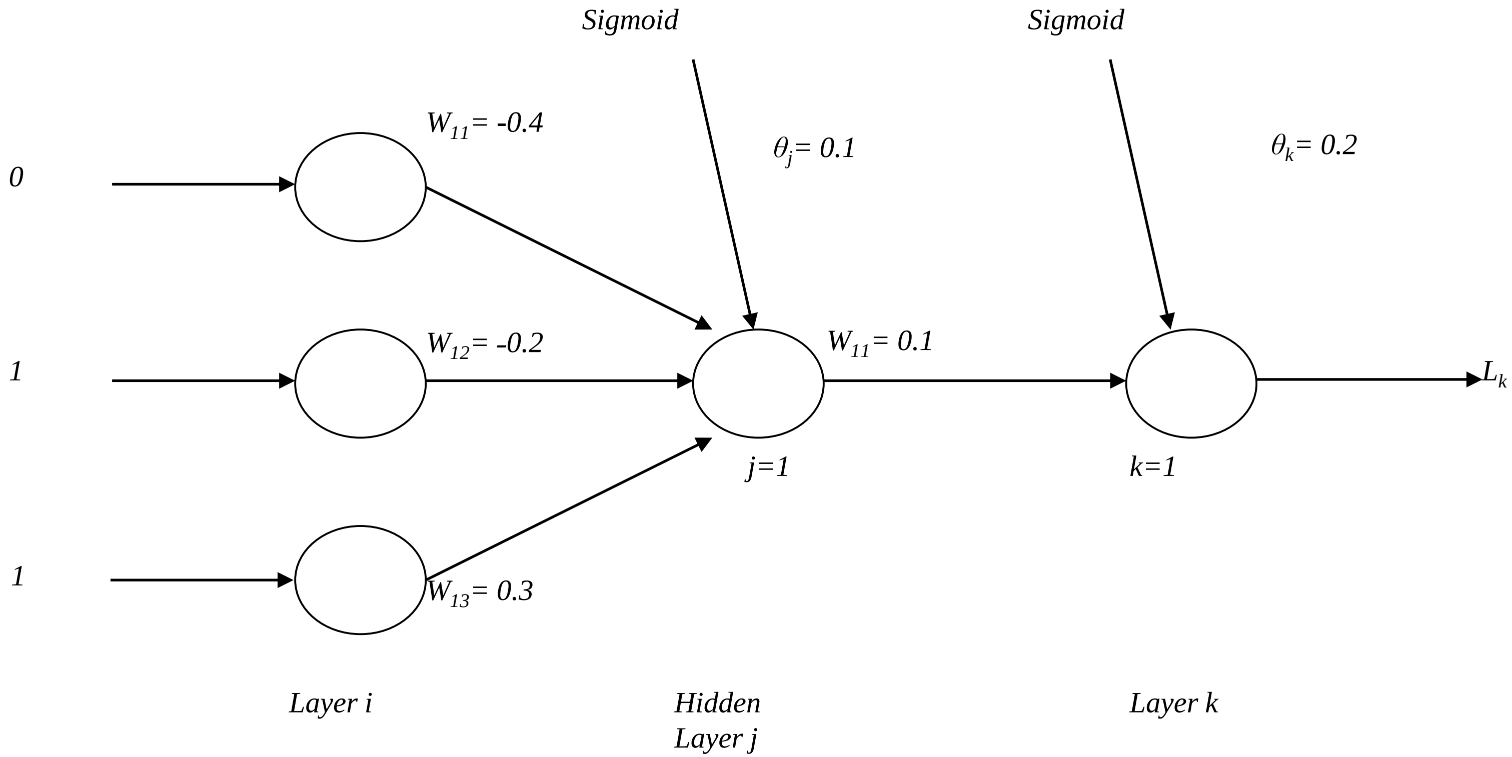


Figure 1: A Multilayer Neural Network

Question Cont'd

i) Calculate the output value of L_k .

ii) Given $t_k = 0$. From the value of L_k, calculate the following values at the first iteration by using Back Propagation algorithm.

- Δw_{11} and w_{11} (new) between output and hidden layer
- Δw_{11}, Δw_{12}, Δw_{13}, *and* w_{11}*(new)*, w_{12}*(new)*, w_{13}*(new)* between hidden layer and input layer.
- Illustrate the new Neural Network.

Question Cont'd

The information for the Neural Networks configurations are as follows.

Given η = 0.4 and α = 0.1.

Back propagation is not required to be derived.

Sigmoid function; $f(x) = (1+e^{-x})^{-1}$

$f(x) = x$

The error signals are as follows.

$\delta_k = L_k\,(1- L_k\,)(\,t_k - L_k\,)$

$\delta_j = Lj\,(1- L_j\,) \sum_k \delta_k\, w_{kj}$

- Adaptions of weights are defined as below.

$\Delta w_{kj}(\,t + 1) = \eta\, \delta_k\, L_j + \alpha \Delta w_{kj}(\,t\,)$

$\Delta w_{ji}(\,t + 1) = \eta\, \delta_j\, L_i + \alpha \Delta w_{ji}(\,t\,)$

[16 Marks]

- **[CO2, P10, C4]**

-

Answer

(i) Calculation of L_k

- Hidden Layer (L_j)

$$y_j(p) = sigmoid\left[\sum_{i=1}^{n} x_i(p) \times w_{ij}(p) - \theta_j\right]$$

$$y_j(p) = sigmoid[(x_1 \times W_{11} + x_2 \times W_{12} + x_3 \times W_{13}) - \theta_j]$$

$$y_j(p) = sigmoid[(0 \times (-0.4) + 1 \times (-0.2) + 1 \times 0.3) - 0.1]$$

$$y_j(p) = sigmoid[0.1 - 0.1]$$

$$y_j(p) = sigmoid[0]$$

$$y_j(p) = L_j = \frac{1}{1 + e^{-(0)}} = 0.50$$

Answer Cont'd

- Output layer

$$y_k(p) = sigmoid\left[\sum_{i=1}^{n} x_i(p) \times w_{ik}(p) - \theta_k\right]$$

$$y_k(p) = sigmoid[(x_1 \times W_{11}) - \theta_k]$$

$$y_k(p) = sigmoid[(0.50 \times 0.1) - 0.2]$$

$$y_k(p) = sigmoid[0.05 - 0.2]$$

$$y_k(p) = sigmoid[-0.15]$$

$$y_k(p) = L_k = 0.46$$

Answer Cont'd

(i) Using back propagation method and information $t_k = 0$ and the L_k obtained from (i).

- Calculation Δw_{11} and w_{11} (new) between output and hidden layer

$$\delta_k = L_k (1 - L_k)(t_k - L_k)$$

$$\delta_k = 0.46 (1 - 0.46)(0 – 0.46) = -0.1143$$

$$\Delta w_{11}(t + 1) = \eta \delta_k L_j + \alpha \Delta w_{11}(t)$$

$$\Delta w_{11} (t + 1) = (0.4)*(-0.1143)*(0.5) + (0.1)*(0) = -0.0229$$

$$w_{11} (t + 1) = w_{11} (t) + \Delta w_{11} (t + 1) = (0.1) + (-0.0229) = 0.0771$$

Answer Cont'd

$\delta j = L_j\,(1 - L_j)\sum \delta_k w_{kj}$

$\delta j = 0.5\,(1 - 0.5)(-0.1143*0.2) = -0.0057$

$\Delta w_{ji}(t+1) = \eta \delta_j L_i + \alpha \Delta w_{ji}(t)$

$\Delta w_{11}(t+1) = \eta \delta_j L_1 + \alpha \Delta w_{11}(t) = 0.4 \times -0.0057 \times 0 + 0.1 \times 0 = 0$

$\Delta w_{12}(t+1) = \eta \delta_j L_2 + \alpha \Delta w_{12}(t) = 0.4 \times -0.0057 \times 1 + 0.1 \times 0 = -2.28 \times 10^{-3}$

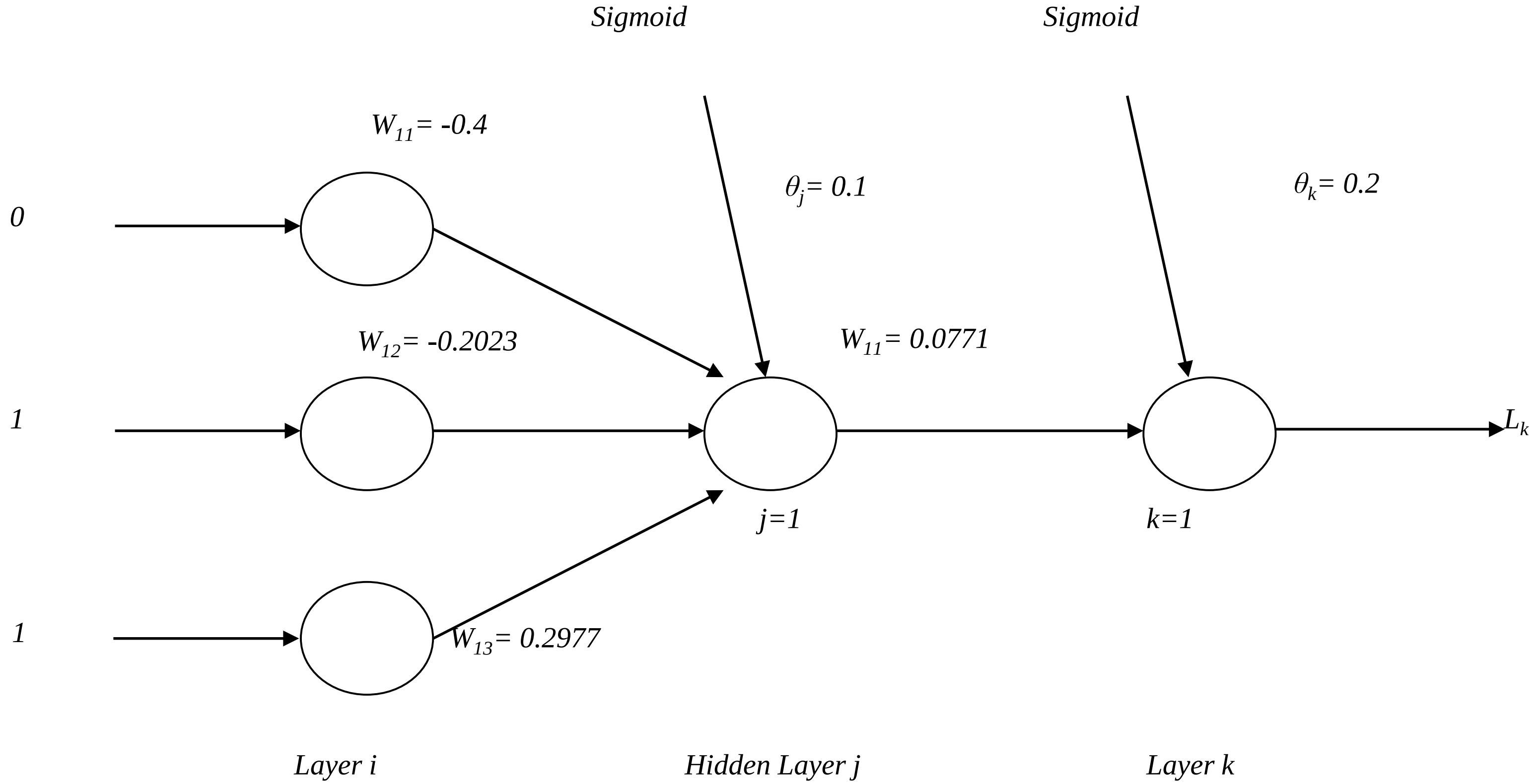
$\Delta w_{13}(t+1) = \eta \delta_j L_3 + \alpha \Delta w_{13}(t) = 0.4 \times -0.0057 \times 1 + 0.1 \times 0 = -2.28 \times 10^{-3}$

$w_{11}(t+1) = w_{11}(t) + \Delta w_{11}(t+1) = -0.4 + 0 = -0.4$

$w_{12}(t+1) = w_{12}(t) + \Delta w_{12}(t+1) = -0.2 - 2.28 \times 10^{-3} = -0.2023$

$w_{13}(t+1) = w_{13}(t) + \Delta w_{12}(t+1) = 0.3 - 2.28 \times 10^{-3} = 0.2977$

New Neural Network



The new neural network

Dr. Nor Maniha Abdul Ghani

Faculty of Electrical and Electronics Engineering
Universiti Malaysia Pahang,
26600, Pekan, Pahang, Malaysia
Phone: +609-424-6087
Fax: +609-424-6000

http://fkee.ump.edu.my/index.php/en/staff-menu/articles-staff/1034-niha-main-profile