MATLAB Deep Learning Notes I

1. Nodes of Neural Network

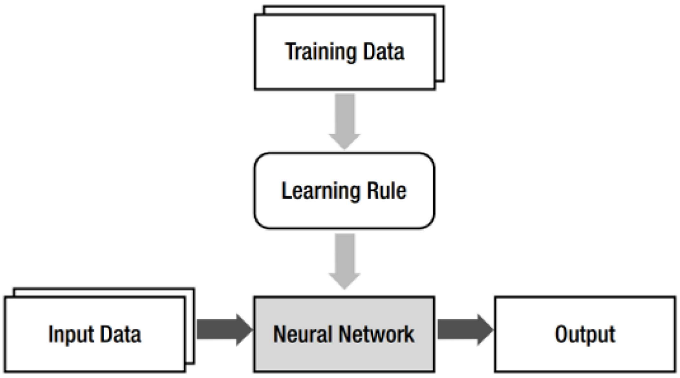

Figure 1: The analogy between the brain and neural network

1. The models of Machine Learning can be implemented in various forms. The neural network is one of them. Fig. 1 illustrates the relationship between Machine Learning and the neural network.

2. The computer stores information at specified locations of the memory, while the brain alters the association of neurons.

3. It is worth pointing out that the neuron itself has no storage capability; it just transmits signals from one neuron to the other. In other words, the brain is gigantic network of these neurons, and the association of the neurons forms specific information.

4. The neural network mimics the neurons association, which is the most important mechanism of the brain, using the weight value. Table.1 summarizes the analogy between the brain and neural network.

Brain	Neural Network
Neuron	Node
Connection of neurons	Connection weight

Table 1: Brain vs. Neural Network.

5. Consider a node that receivers three inputs as shown below:

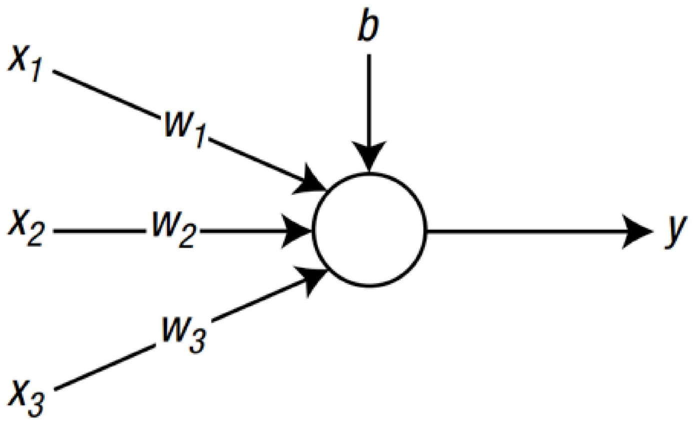

Figure 2: A node that receives three inputs

Then we have

$$v = (w_1 \times x_1) + (w_2 \times x_2) + (w_3 \times x_3) + b \tag{1}$$

where w_1, w_2, and w_3 denote the weights, x_1, x_2, and x_3 are the input signals, and b refers to bias. Equation (1) indicates that the signal with a greater weight has a greater effect. Then lets rewrite this equation, so we have

$$v = \mathbf{w}\mathbf{x} + b \tag{2}$$

where $\mathbf{w} = \begin{bmatrix} w_1 & w_2 & w_3 \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$, then consider the sum v as the input of the activation function $\varphi(\cdot)$, so we have the final output

$$y = \varphi(v) = \varphi(\mathbf{w}\mathbf{x} + b) \tag{3}$$

6. We will introduce several different activation functions later.

2. Layer of Neural Network

1. A variety of neural networks can be created depending on how the nodes are connected. One of the examples as shown below

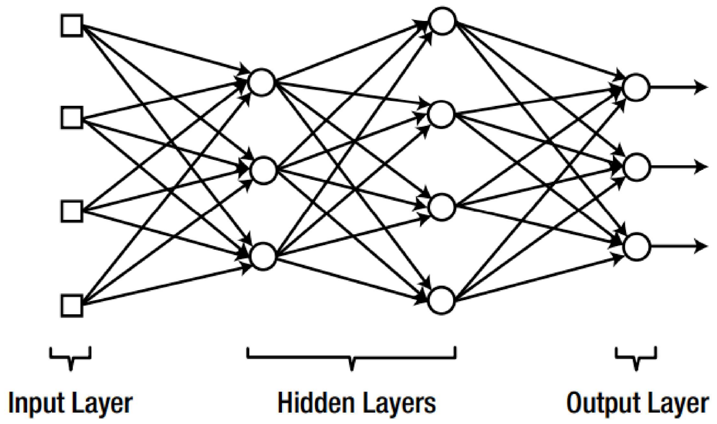

Figure 3: A layered structure of nodes

2. The layers in between the input and output layers are called hidden layers. They are given this name because they are not accessible from the outside of the neural network (NN). The neural network that has a single hidden layer is called a shallow neural network or a vanilla neural network. A multi-layer neural network that contains two or more hidden layers is called a deep neural network.

Single-Layer Neural Network		Input Layer – Output Layer
Multi-Layer Neural Network	Shallow Neural Network	Input Layer – Hidden Layer – Output Layer
	Deep Neural Network	Input Layer – Hidden Layers – Output Layers

Figure 4: Single-Layer Network vs. Multi-Layer Neural Network

3. Supervised Learning of a Neural Network

Algorithm 1 Basic Supervised Learning of a Neural Network

Input:

Training Data which is formatted as {input, correct output};

Output:

The neural network with appropriate weights;

1: Initialize the weights at adequate values;
2: **for** each $i \in [1, n]$ **do**
3: Take the Training Data into the NN;
4: Obtain the output from NN;
5: Calculate the error from the correct input;
6: Adjust the weights to reduce the error;
7: **end for**

The training of supervised learning is a process that modifies the model to reduce the difference between the correct output and models output. The only difference is that the modification of the model becomes the changes in weights for the neural network.

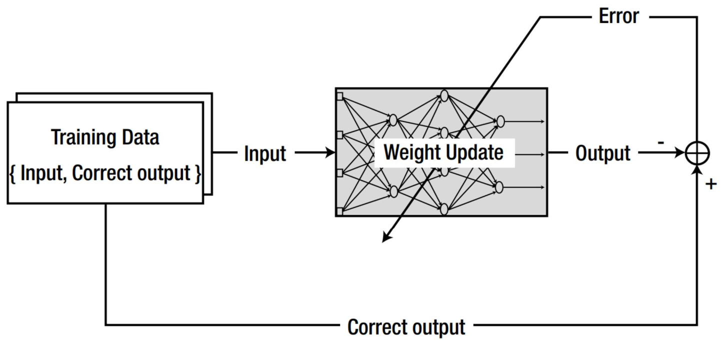

Figure 5: The concept of supervised learning

4. Training of a Single-Layer Neural Network: Delta Rule

1. In order to train the neural network with new information, the weights should be changed accordingly.

2. The systematic approach to modifying the weights according to the given information is called the learning rule, which is a vital component in neural network research.

3. We consider the Delta rule(also called Adaline rule or Widrow-Hoff rule) first, although it is just for single-layer neural network.

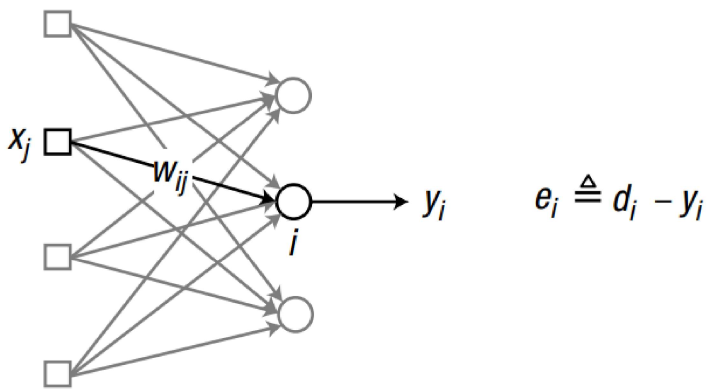

Figure 6: A single-layer neural network

4. If an input node contributes to the error of the output node, the weight between the two nodes is adjusted in proportion to the input value, x_j and the output error, e_i.

5.This rule can be expressed in equation as:

$$w_{ij} \leftarrow w_{ij} + \alpha e_i x_j \tag{4}$$

where x_j is the output from the input node j, $(j = 1, 2, 3)$; e_i denotes the error of the output node j; w_{ij} represents the weight between the output node i and input node j; w_{ij} is the weight between the output node i and input node j; α refers to learning rate $(0 < \alpha \leq 1)$.

6.The learning rate α determines how much the weight is changed per time. If α is too high, the output wanders around the solution and fails to converge. In contrast, if it is too low, the calculation reaches the solution too slowly.

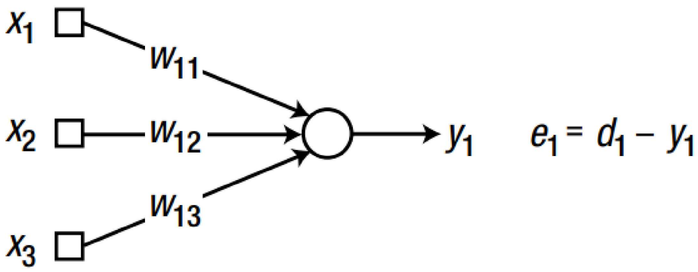

Figure 7: A single-layer neural network with three input nodes and one output node

7. From Fig.7, we have

$$w_{11} \leftarrow w_{11} + \alpha e_1 x_1$$

$$w_{12} \leftarrow w_{12} + \alpha e_1 x_2$$

$$w_{13} \leftarrow w_{13} + \alpha e_1 x_3$$

so the training process using delta rule for the single-layer neural network is

Algorithm 2 The Training Process using Delta Rule for the Single-Layer Neural Network

Input:
Training Data which is formatted as {input, correct output};
Output:
The neural network with appropriate weights;

1: Initialize α and the weights at adequate values;
2: **for** epoch $i \in [1, \text{the number of epoch}]$ **do**
3: **for** each $i \in [1, 3]$ **do**
4: Take the Training Data into the NN;
5: Obtain the output y_1 from NN;
6: Calculate the error from the correct input($e_1 = d_1 - y_1$);
7: Adjust the weights to reduce the error($w_{1i} \leftarrow w_{1i} + \alpha e_1 x_i$);
8: **end for**
9: **end for**

8. The number of training iterations, in each of which all training data goes through once, is called an epoch.

9. Although the equations may vary depending on the learning rule, the essential concepts are relatively the same as Fig.8

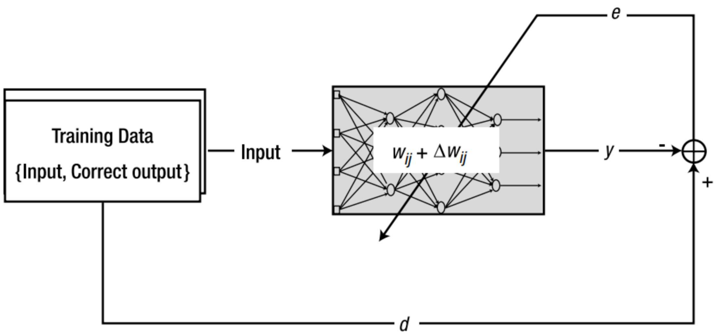

Figure 8: Training Process