Dement Neurocognitive Disord. 2018 Sep;17(3):83-89
https://doi.org/10.12779/dnd.2018.17.3.83
pISSN 1738-1495·eISSN 2384-0757

DND Dementia and Neurocognitive Disorder

Review Article

Check for updates

Artificial Neural Network: Understanding the Basic Concepts without Mathematics

Su-Hyun Han,[1] **Ko Woon Kim**,[2] **SangYun Kim**,[3] **Young Chul Youn**[1]

[1]Department of Neurology, Chung-Ang University College of Medicine, Seoul, Korea
[2]Department of Neurology, Chonbuk National University Hospital, Jeonju, Korea
[3]Department of Neurology, Seoul National University College of Medicine and Seoul National University Bundang Hospital, Seongnam, Korea

OPEN ACCESS

Received: Sep 30, 2018
Revised: Nov 4, 2018
Accepted: Nov 20, 2018

Correspondence to
Young Chul Youn, MD, PhD
Department of Neurology, Chung-Ang University College of Medicine, Chung-Ang University Hospital, 102 Heukseok-ro, Dongjak-gu, Seoul 06973, Korea.
E-mail: neudoc@cau.ac.kr

© 2018 Korean Dementia Association
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (https://creativecommons.org/licenses/by-nc/4.0/) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

ORCID iDs
Su-Hyun Han
https://orcid.org/0000-0002-3084-6985
SangYun Kim
https://orcid.org/0000-0002-9101-5704
Young Chul Youn
https://orcid.org/0000-0002-2742-1759

Funding
This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2017S1A6A3A01078538).

Conflict of Interest
The authors have no financial conflicts of interest.

ABSTRACT

Machine learning is where a machine (i.e., computer) determines for itself how input data is processed and predicts outcomes when provided with new data. An artificial neural network is a machine learning algorithm based on the concept of a human neuron. The purpose of this review is to explain the fundamental concepts of artificial neural networks.

Keywords: Machine Learning; Artificial Intelligence; Neural Networks; Deep Learning

WHAT IS THE DIFFERENCE BETWEEN MACHINE LEARNING AND A COMPUTER PROGRAM?

A computer program takes input data, processes the data, and outputs a result. A programmer stipulates how the input data should be processed. In machine learning, input and output data is provided, and the machine determines the process by which the given input produces the given output data.[1-3] This process can then predict the unknown output when new input data is provided.[1-3]

So, how does machine determine the process? Consider a linear model, $y = Wx + b$. If we are given the x and y values shown in **Table 1**, we know intuitively that $W = 2$ and $b = 0$.

However, how does a computer know what W and b are? W and b are randomly generated initially. For example, if we initially say $W = 1$ and $b = 0$ and input the x values from **Table 1**, the

Table 1. Example dataset for the linear model $y = Wx + b$

x	y
1	2
2	4
3	6
4	8
5	10
10	20
20	40
30	60

Generated by KOMET PRESS

Author Contributions

Conceptualization: Kim SY; Supervision: Kim KW, Youn YC; Writing - original draft: Han SH, Kim KW; Writing - review & editing: Youn YC.

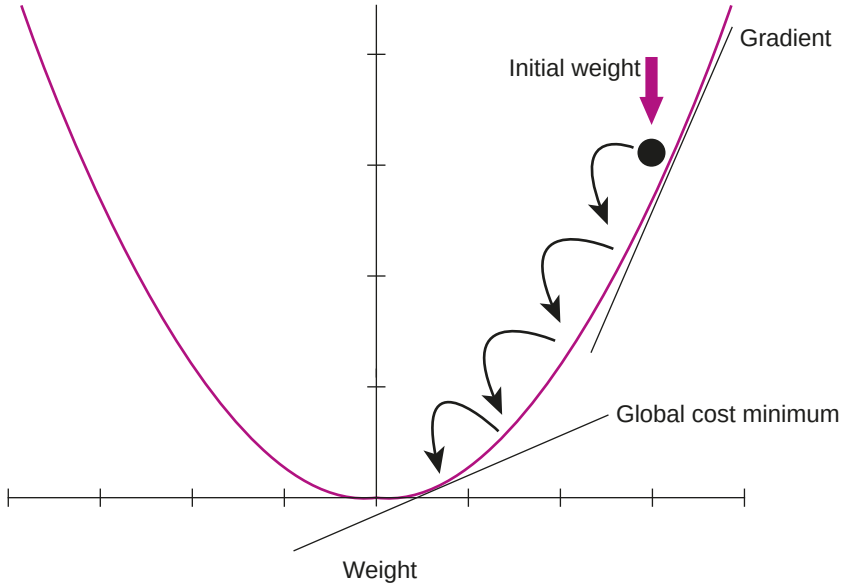

Fig. 1. A graph of a cost function (modified from https://rasbt.github.io/mlxtend/user_guide/general_concepts/gradient-optimization/).

predicted *y* values do not match the *y* values in the **Table 1**. At this point, there is a difference between the correct *y* values and the predicted *y* values; the machine gradually adjusts the values of *W* and *b* to reduce this difference. The difference between the predicted values and the correct values is called the cost function. Minimizing the cost function makes predictions closer to the correct answers.[4-7]

The costs that correspond to given *W* and *b* values are shown in the **Fig. 1** (obtained from https://rasbt.github.io/mlxtend/user_guide/general_concepts/gradient-optimization/).[8-10] Given the current values of *W* and *b*, the gradient is obtained. If the gradient is positive, the values of *W* and *b* are decreased. If the gradient is negative, the values of *W* and *b* are increased. In other words, the values of *W* and *b* determined by the derivative of the cost function.[8-10]

ARTIFICIAL NEURAL NETWORK

The basic unit by which the brain works is a neuron. Neurons transmit electrical signals (action potentials) from one end to the other.[11] That is, electrical signals are transmitted from the dendrites to the axon terminals through the axon body. In this way, the electrical signals continue to be transmitted across the synapse from one neuron to another. The human brain has approximately 100 billion neurons.[11,12] It is difficult to imitate this level of complexity with existing computers. However, Drosophila have approximately 100,000 neurons, and they are able to find food, avoid danger, survive, and reproduce.[13] Nematodes have 302 neurons and they survive well.[14] This is a level of complexity that can be replicated well even with today's computers. However, nematodes can perform much better than our computers.

Let's think of the operative principles of neurons. Neurons receive signals and generate other signals. That is, they receive input data, perform some processing, and give an output.[11,12] However, the output is not given at a constant rate; the output is generated when the input

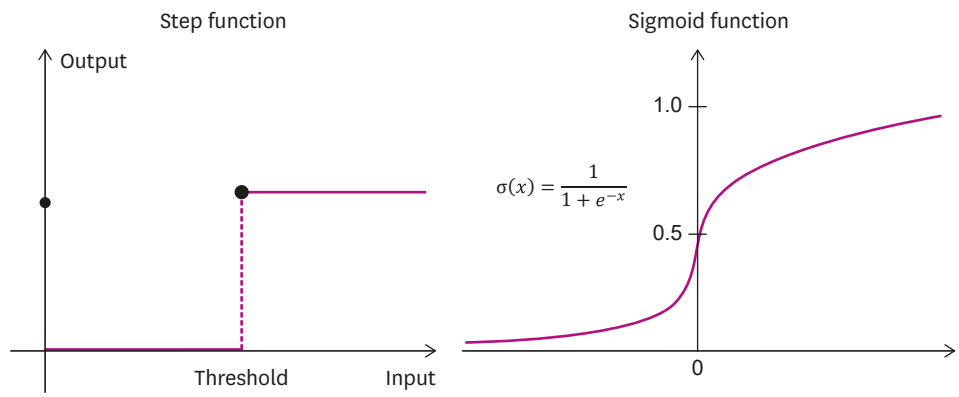

Fig. 2. A step function and a sigmoid function.

exceeds a certain threshold.[15] The function that receives an input signal and produces an output signal after a certain threshold value is called an activation function.[11,15] As shown in **Fig. 2**, when the input value is small, the output value is 0, and once the input value rises above the threshold value, a non-zero output value suddenly appears. Thus, the responses of biological neurons and artificial neurons (nodes) are similar. However, in reality, artificial neural networks use various functions other than activation functions, and most of them use sigmoid functions.[4,6] which are also called logistic functions. The sigmoid function has the advantage that it is very simple to calculate compared to other functions. Currently, artificial neural networks predominantly use a weight modification method in the learning process.[4,5,7] In the course of modifying the weights, the entire layer requires an activation function that can be differentiated. This is because the step function cannot be used as it is. The sigmoid function is expressed as the following equation.[6,16]

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

Biological neurons receive multiple inputs from pre-synaptic neurons.[11] Neurons in artificial neural networks (nodes) also receive multiple inputs, then they add them and process the sum with a sigmoid function.[5,7] The value processed by the sigmoid function then becomes the output value. As shown in **Fig. 3**, if the sum of inputs A, B, and C exceeds the threshold and the sigmoid function works, this neuron generates an output value.[4]

A neuron receives input from multiple neurons and transmits signals to multiple neurons.[4,5]

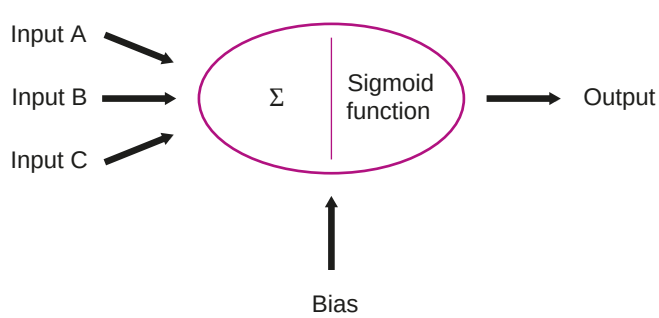

Fig. 3. Input and output of information from neurons.

https://doi.org/10.12779/dnd.2018.17.3.83