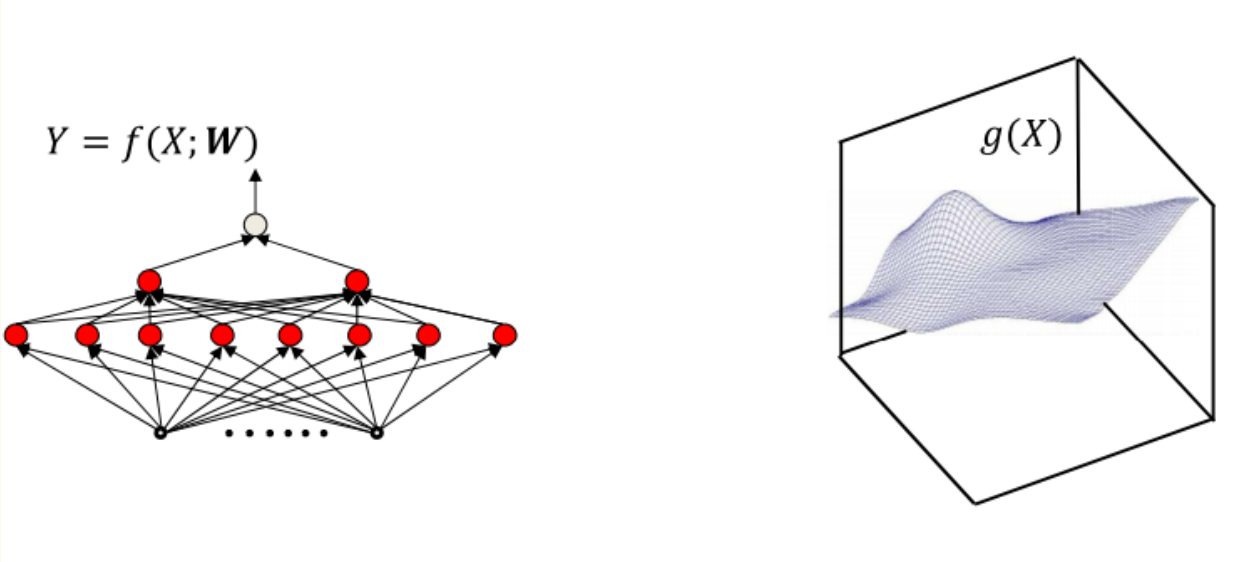
# LEARNING A NEURAL NET

—



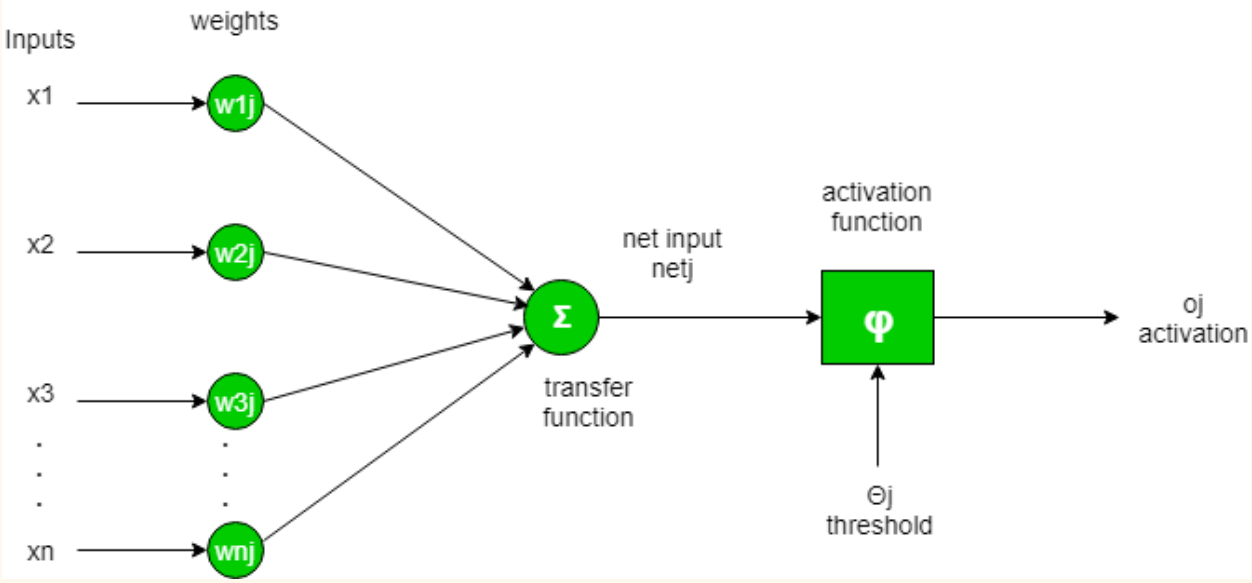

"Learning" a neural network = determining the parameters of the network (weights and biases) required for it to model a desired function – The network must have sufficient capacity to model the function

• Ideally, we would like to optimize the network to represent the desired function everywhere

• However this requires knowledge of the function everywhere

• Instead, we draw "input-output" training instances from the function and estimate network parameters to "fit" the input-output relation at these instances.

# Activation Functions

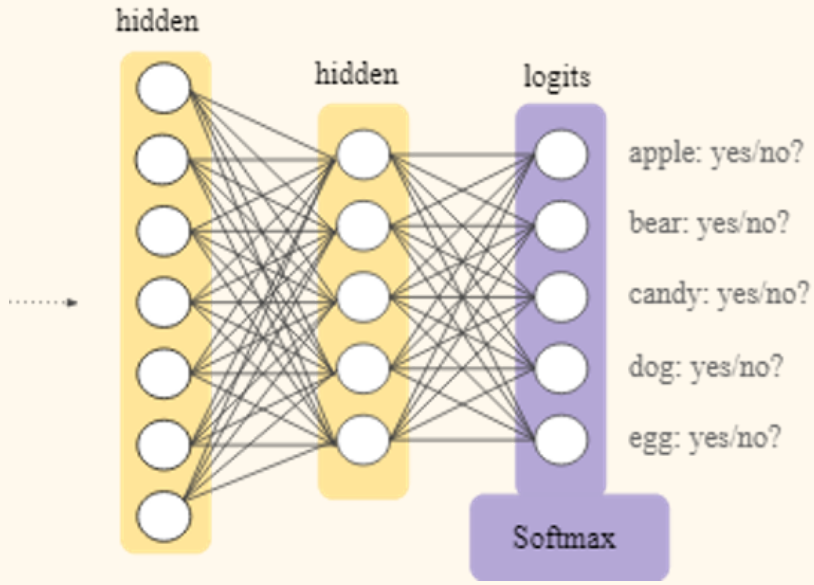



In a **neural network**, the **activation function** is responsible for transforming the summed weighted input from the node into the **activation** of the node or output for that input.

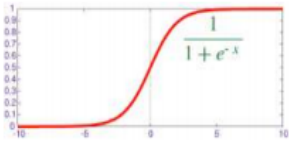
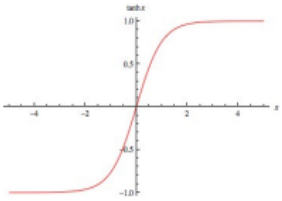
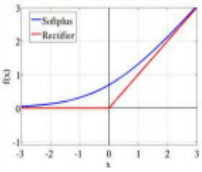
TYPES -

- ❏ Softmax
- ❏ Sigmoid
- ❏ Tanh
- ❏ ReLU

Softmax -




Other Activation functions -

| | |
|---|---|
|  $f(z) = \frac{1}{1+\exp(-z)}$ | $f'(z) = f(z)(1-f(z))$ |
|  $f(z) = \tanh(z)$ | $f'(z) = (1-f^2(z))$ |
|  $f(z) = \begin{cases} z, & z \geq 0 \\ 0, & z < 0 \end{cases}$ | [*] $f'(z) = \begin{cases} 1, z \geq 0 \\ 0, z < 0 \end{cases}$ |
| $f(z) = \log(1+\exp(z))$ | $f'(z) = \frac{1}{1+\exp(-z)}$ |

FORWARD AND BACKWARD PROPAGATION

Before we start let’s have a look at a sample data set which will be used to train the neural network-

| gender | age | hypertension | does_smoke | stroke |
|---|---|---|---|---|
| 1 | 3 | 0 | 0 | 0 |
| 1 | 58 | 1 | 1 | 1 |
| 0 | 8 | 0 | 0 | 0 |
| 0 | 70 | 0 | 1 | 1 |
| 1 | 14 | 0 | 0 | 0 |
| 0 | 47 | 0 | 0 | 0 |
| 0 | 52 | 0 | 1 | 1 |
| 0 | 75 | 0 | 0 | 0 |
| 0 | 32 | 0 | 1 | 0 |