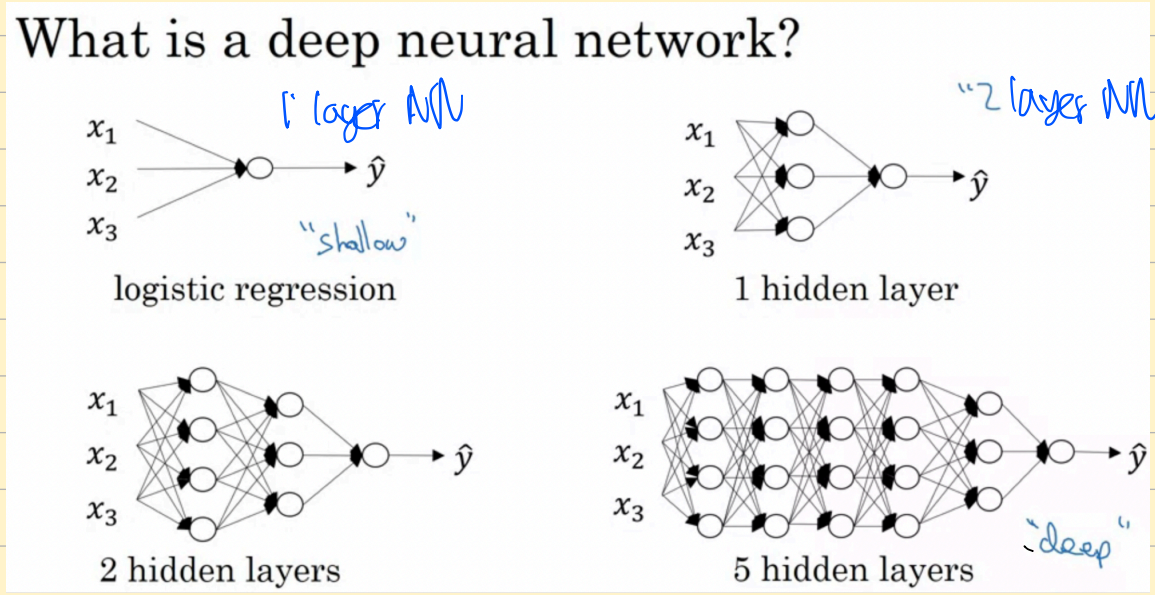
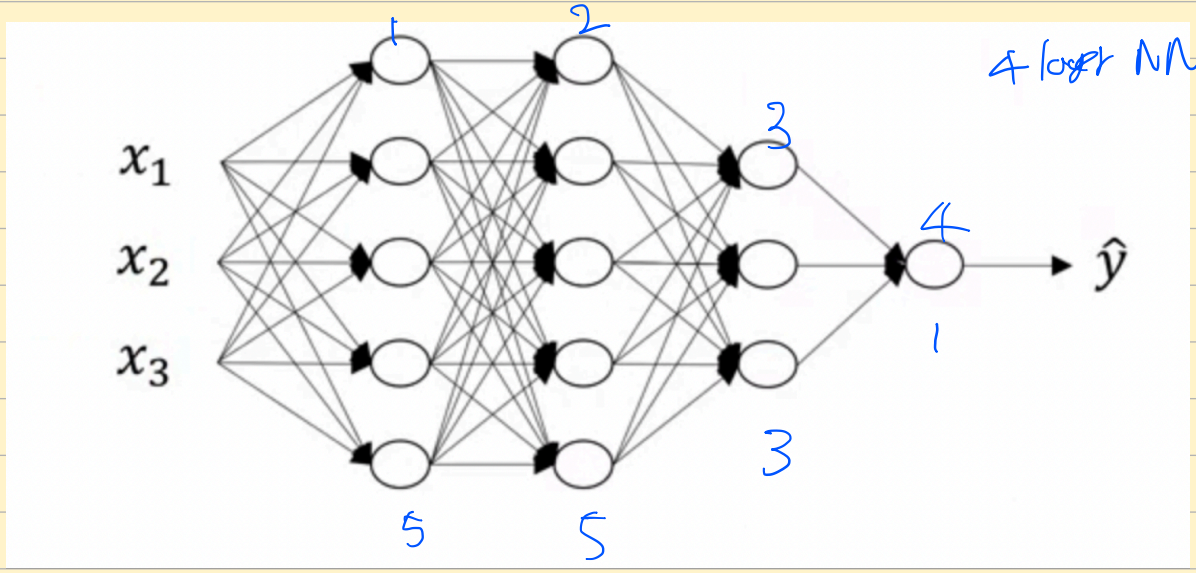
Deep neural network



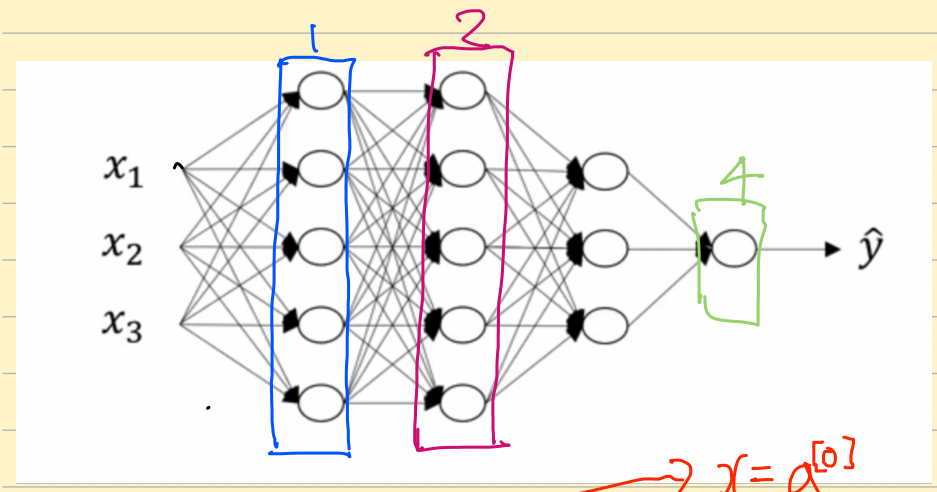
- 일반적으로 hidden layer의 개수가 2개 이상인 network부터 Deep Neural Network라고 한다.



* notation.

- L : layer의 개수, hidden layer와 output layer 개수의 합. 위의 경우 $L=4$
- $n^{[l]}$: layer l에 있는 unit 수. $n^{[1]}=5$, $n^{[2]}=5$ ...
- $a^{[l]}$: layer l의 activation 들.

Forward propagation in a deep network

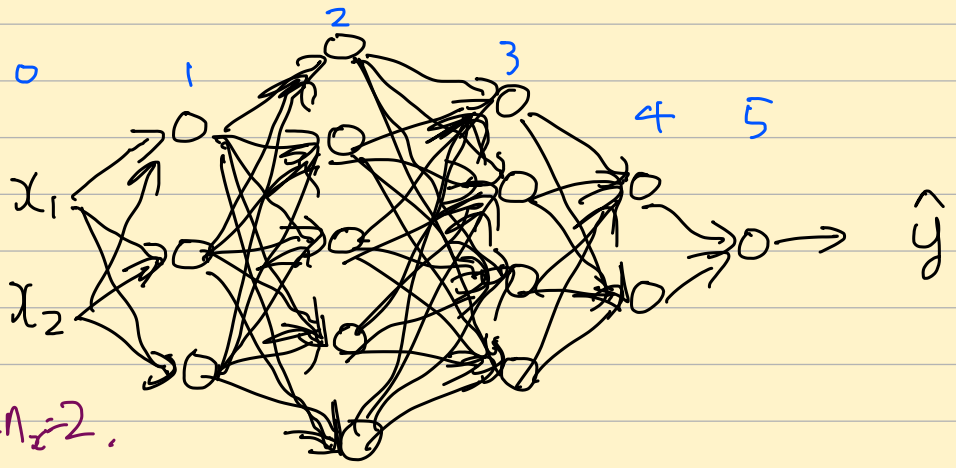


$x = a^{[0]}$

$$z^{[1]} = W^{[1]} x + b^{[1]}$$
$$a^{[1]} = g^{[1]}(z^{[1]})$$
$$z^{[2]} = W^{[2]} a^{[1]} + b^{[2]}$$
$$a^{[2]} = g^{[2]}(z^{[2]})$$
$$z^{[3]} = W^{[3]} a^{[2]} + b^{[3]}$$
$$a^{[3]} = g^{[3]}(z^{[3]})$$
$$z^{[4]} = W^{[4]} a^{[3]} + b^{[4]}$$
$$a^{[4]} = g^{[4]}(z^{[4]}) = \hat{y}$$

- linear transformation과 non-linear transformation을 반복적으로 수행하는 것이 Forward Propagation!
- x는 example 하나에 대한 표현. 대문자 X는 example 여러개를 합친것
- input layer를 $a^{[0]}$으로 표기하는 이유는 activation function을 맞춰서 통일성을 위함.

- Parameters $W^{[l]}$ and $b^{[l]}$



$n^{[1]} = 3$ $n^{[2]} = 5$ $n^{[3]} = 4$ $n^{[4]} = 2$ $n^{[5]} = 1$

$$z^{[1]} = W^{[1]} \cdot a^{[0]} + b^{[1]}$$

(3,1) (3,2) (2,1) (3,1)

$(n^{[1]}, 1)$ $(n^{[1]}, n^{[0]})$ $(n^{[0]}, 1)$ $(n^{[1]}, 1)$

$$z^{[2]} = W^{[2]} a^{[1]} + b^{[2]}$$

(5,1) (5,3) (3,1) (5,1)

$(n^{[2]}, 1)$ $(n^{[2]}, n^{[1]})$ $(n^{[1]}, 1)$ $(n^{[2]}, 1)$

일반화하면 $W^{[l]} : (n^{[l]}, n^{[l-1]})$

$b^{[l]} : (n^{[l]}, 1)$

$dW^{[l]} : (n^{[l]}, n^{[l-1]})$

$db^{[l]} : (n^{[l]}, 1)$

Why deep representations?

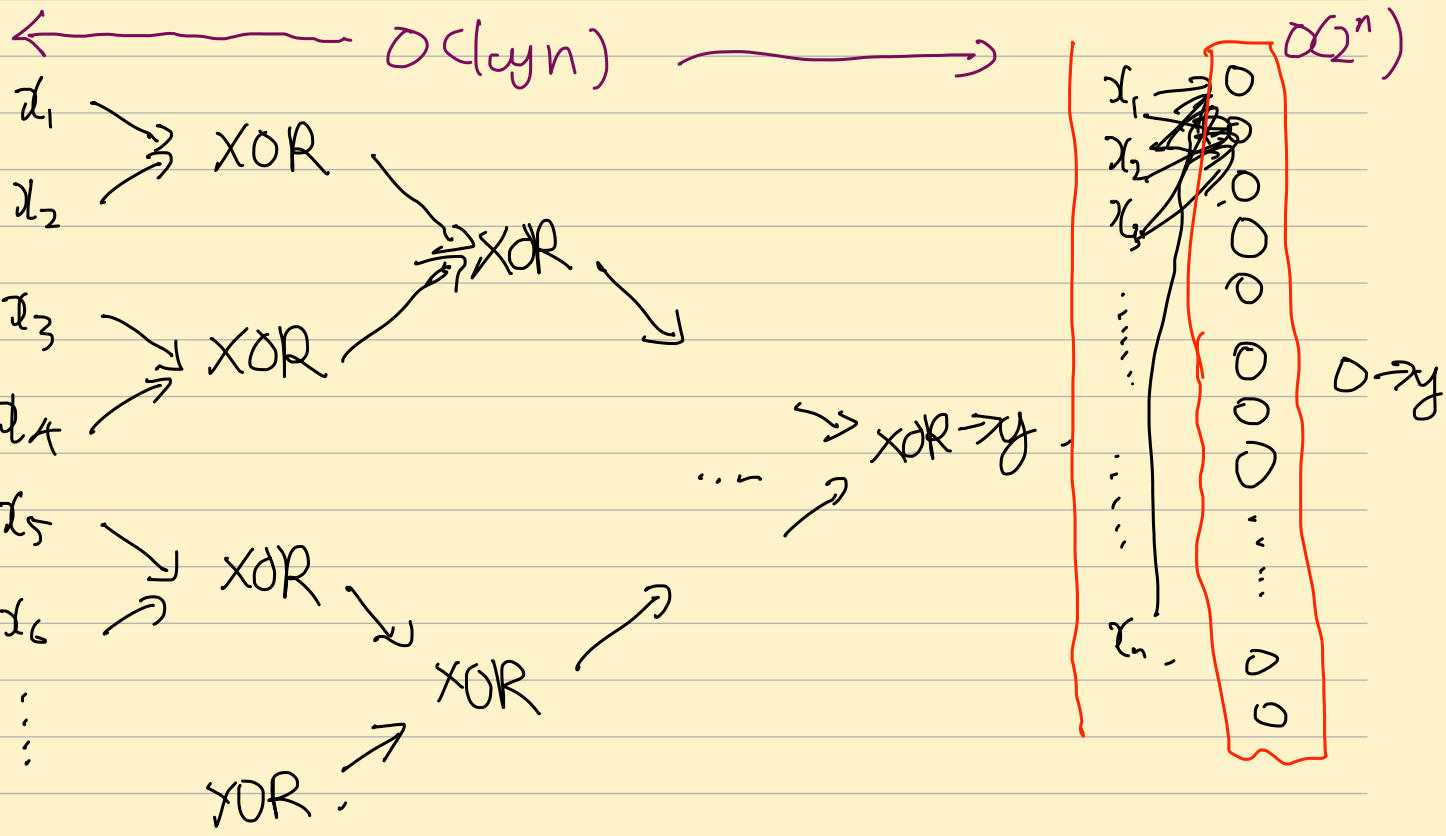
- Deep Neural Network는 데이터의 관계를 단순한 것에서 복잡한 것으로 만든다.

ex) 1. Face Recognition,
   - Image → Edges → Face Parts → Faces
2. Audio Recognition
   - Audio → Low level sound → Phonemes → word → sentence

- Circuit Theory and Deep Learning.

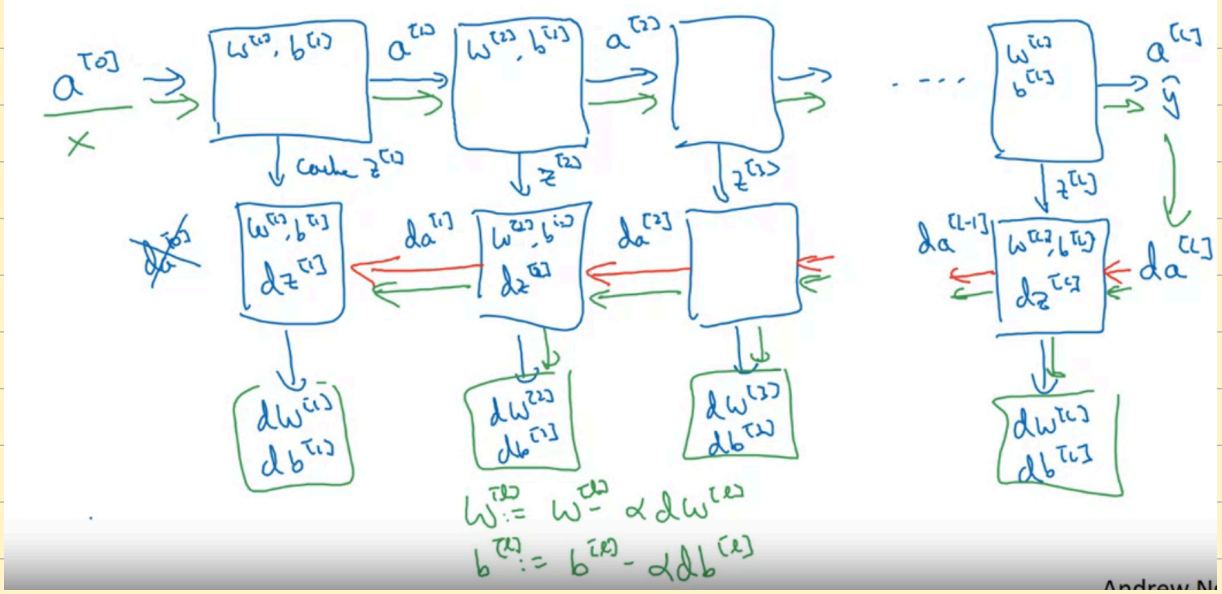
⇒ Shallow Network를 사용하게 되면, 계산에 필요한 hidden unit의 수가 기하급수적으로 늘어나게 된다.

$y = x_1 \text{ XOR } x_2 \text{ XOR } x_3 \text{ XOR } \dots \text{ XOR } x_n$.



Building blocks of deep nn

Forward and backward functions.



Parameters vs Hyperparameters

Parameter : W와 b가 있다.

Hyper Parameter : 학습하기 전에 결정해야 하는 parameters.

ex) learning rate, gradient descent iteration 회수,
hidden layer L의 개수, hidden unit의 수 n,
Activation function.

What does this have to do with the brain
The analogy that "It is like the brain" has become really an oversimplified explanation.
There is a very simplistic analogy between a single logistic unit and a single neuron in the brain.
No human today understand how a human brain neuron works.
No human today know exactly how many neurons on the brain.
Deep learning in Andrew's opinion is very good at learning very flexible, complex functions to learn X to Y mappings, to learn input-output mappings (supervised learning).
The field of computer vision has taken a bit more inspiration from the human brains then other disciplines that also apply deep learning.
NN is a small representation of how brain work. The most near model of human brain is in the computer vision (CNN)