

CHƯƠNG 2

MẠNG NEURAL CƠ BẢN (P2)

Khoa Khoa học và Kỹ thuật thông tin
Bộ môn Khoa học dữ liệu

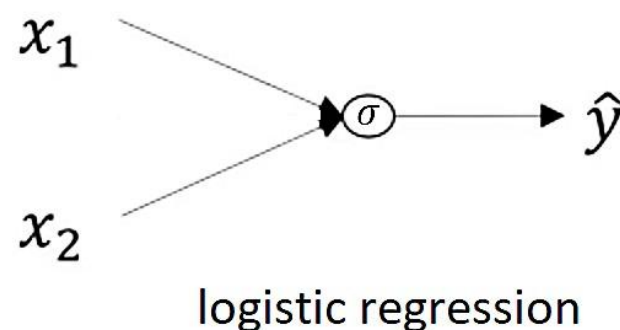
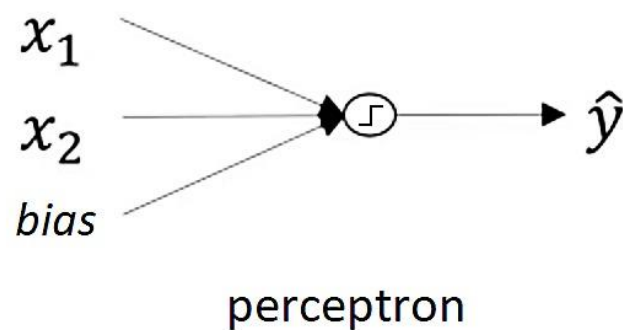
NỘI DUNG

1. Mạng neural nhiều lớp.
2. Truyền xuôi.
3. Chiều của ma trận trong biểu diễn.
4. Truyền ngược.
5. Các loại tham số.

Mạng neural nhiều lớp

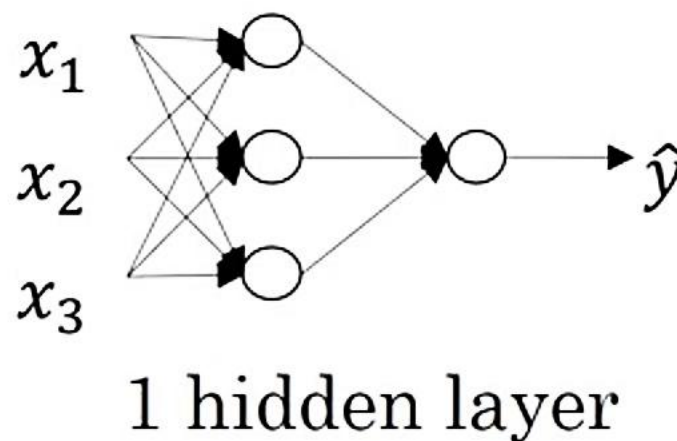
Dẫn nhập

- Trong các mô hình máy học truyền thống như Logistic regression, Perceptron, các thông tin từ dữ liệu chỉ đi qua 1 hàm duy nhất.



Mạng neural sâu

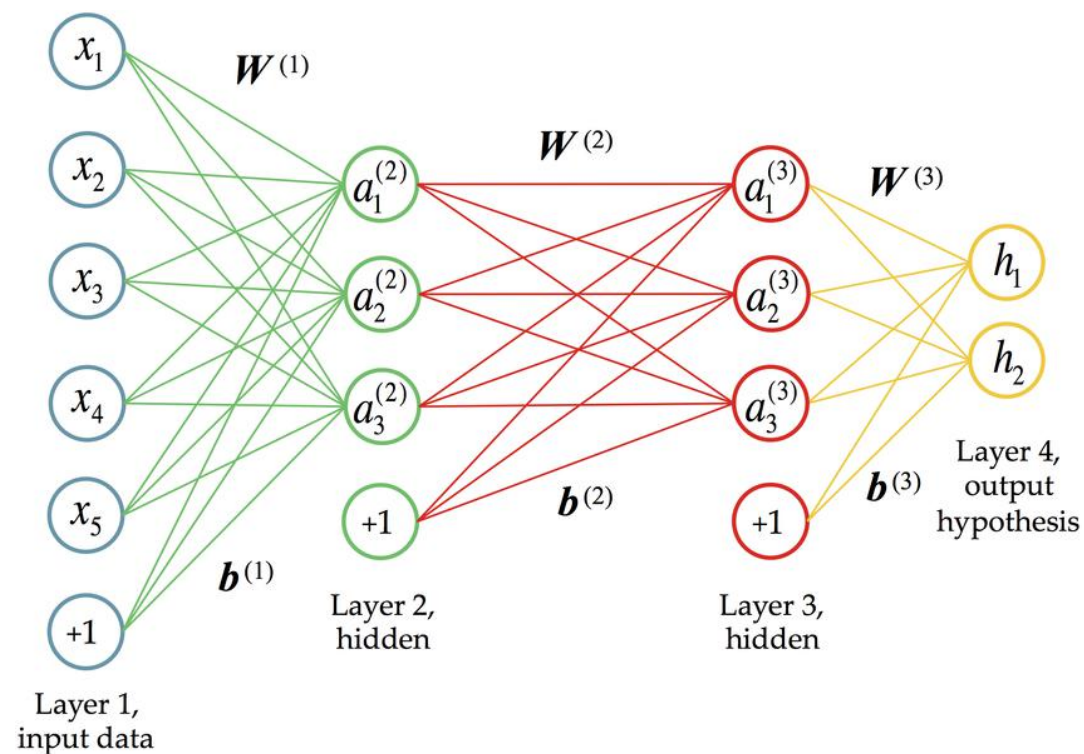
- Ở bài trước, ta đã xem xét về mạng neural 1 lớp ẩn như sau:



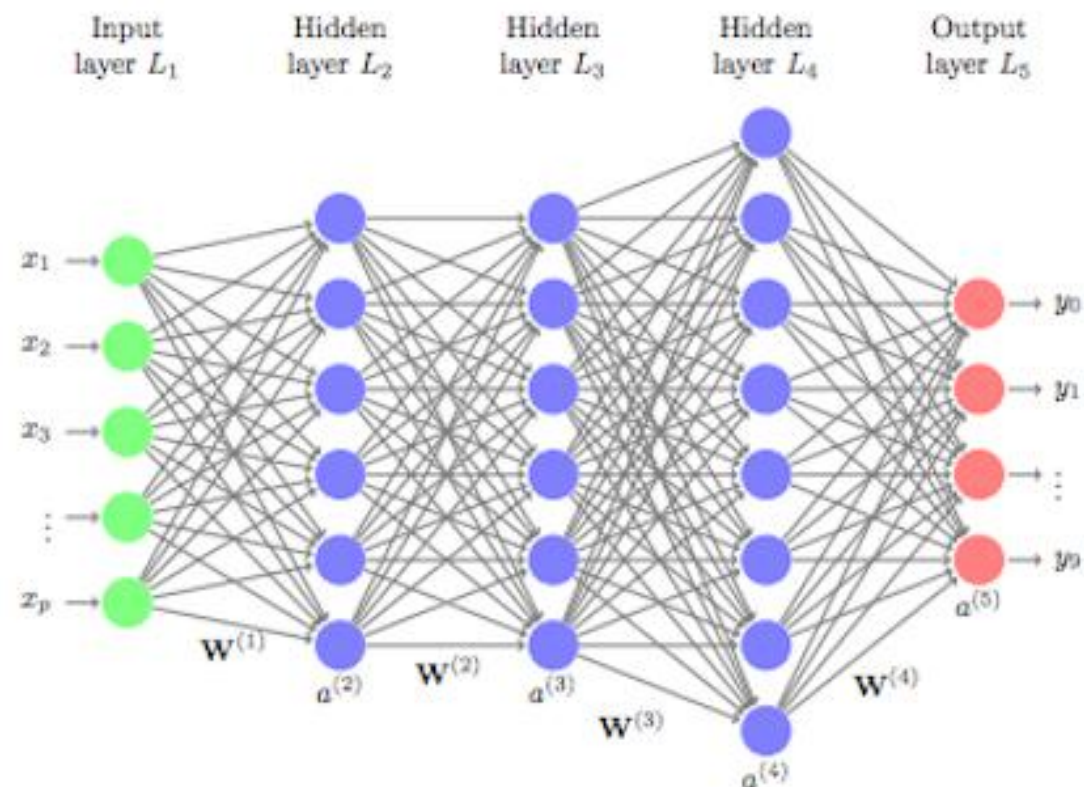
- Trên thực tế, ta có thể thêm các lớp ẩn (hidden layer) vào mạng neural. **Càng nhiều lớp ẩn, mạng neural càng sâu (deep).**

Ví dụ

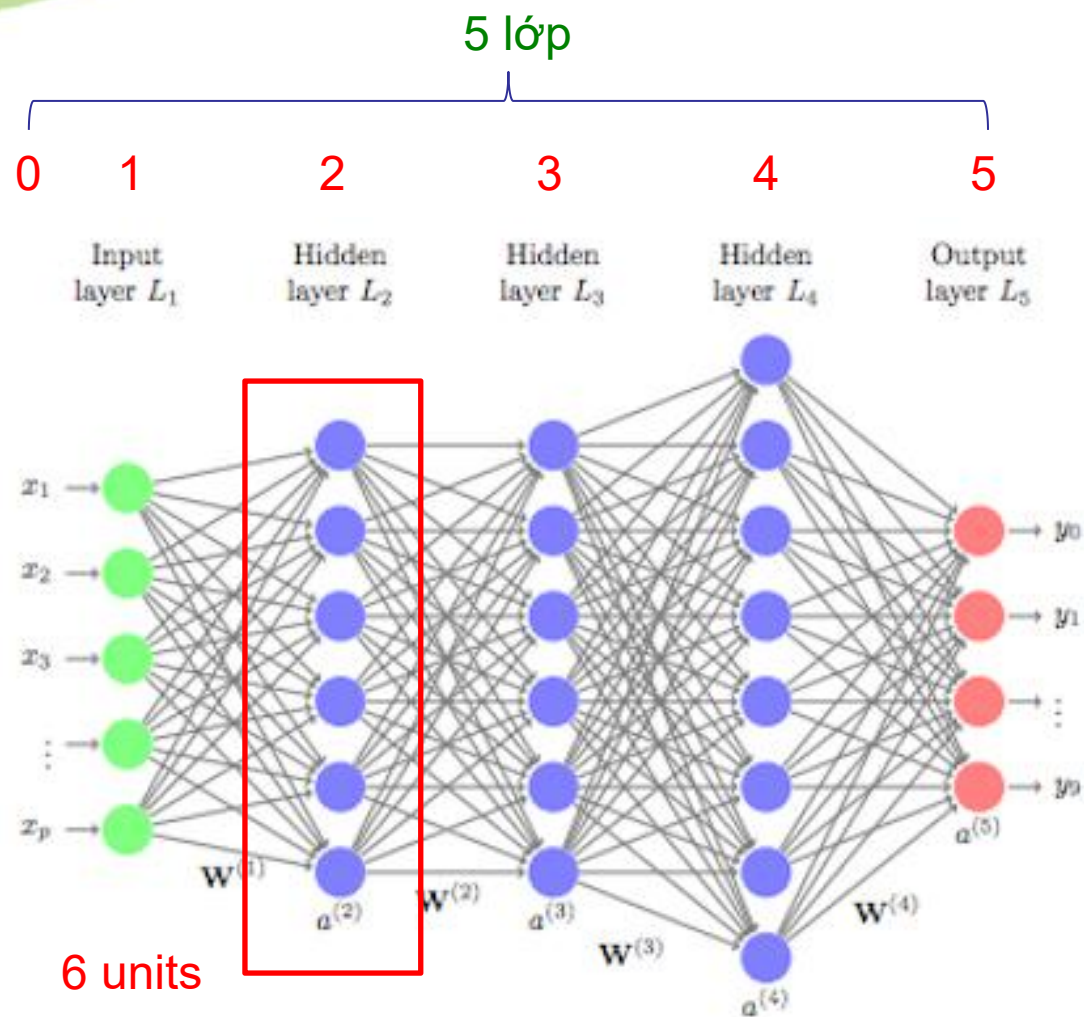
2 lớp ẩn



3 lớp ẩn



Các ký hiệu

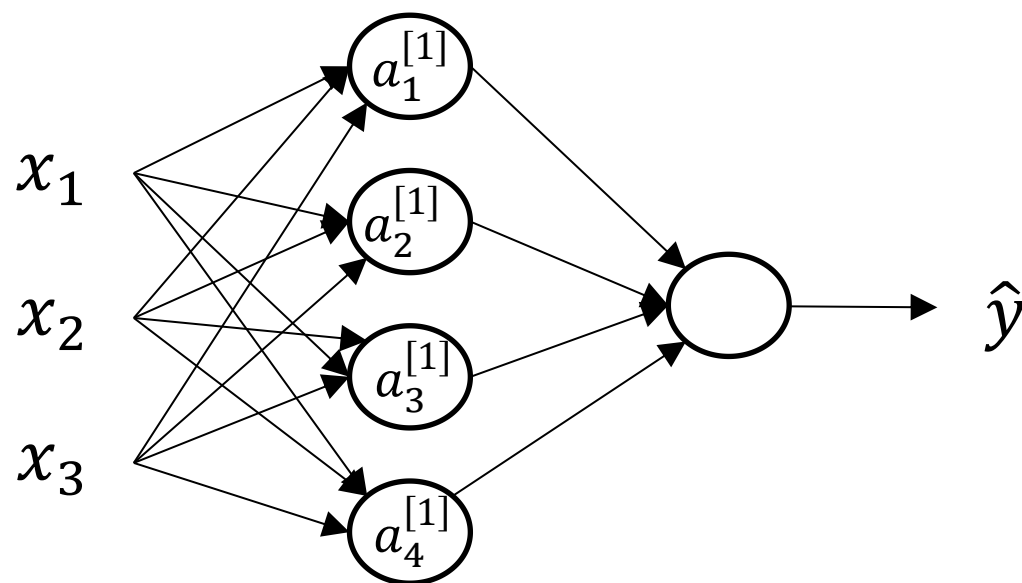


- Các lớp được đánh thứ tự từ 0 đến n , với n là số lượng lớp.
- Dữ liệu input vào được xem là lớp 0.
- Input layer và output layer cũng được tính là một lớp.
- Input layer được tính là lớp thứ 1, ký hiệu là $L^{(1)}$.
- Mỗi layer sẽ có số lượng neural (unit) nhất định.

TRUYỀN XUÔI

MẠNG 1 LỚP ẨN

Given input x :



Truyền xuôi
(feed forward)

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

(4x1) (4x3) (3x1) (4x1)

$$a^{[1]} = \sigma(z^{[1]})$$

(4x1) (4x1)

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

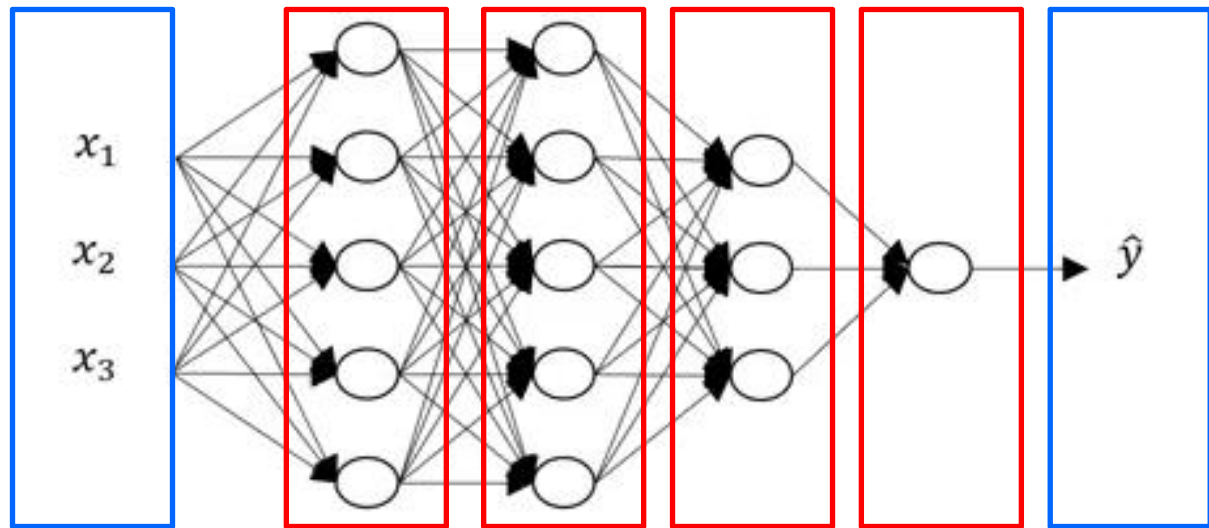
(1x1) (1x4) (4x1) (1x1)

$$a^{[2]} = \sigma(z^{[2]})$$

(1x1) (1x1)

MẠNG 3 LỚP ẨN

$X = a^{[0]}$ $z^{[1]}$ $z^{[2]}$ $z^{[3]}$ $z^{[4]}$
 $a^{[1]}$ $a^{[2]}$ $a^{[3]}$ $a^{[4]}$ $\hat{y} = g^{[4]}(z^{[4]}) = a^{[4]}$



Lớp 1:

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

$$a^{[1]} = g^{[1]}(z^{[1]})$$

Lớp 2:

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

$$a^{[2]} = g^{[2]}(z^{[2]})$$

Lớp 3:

$$z^{[3]} = W^{[3]} * a^{[2]} + b^{[3]}$$

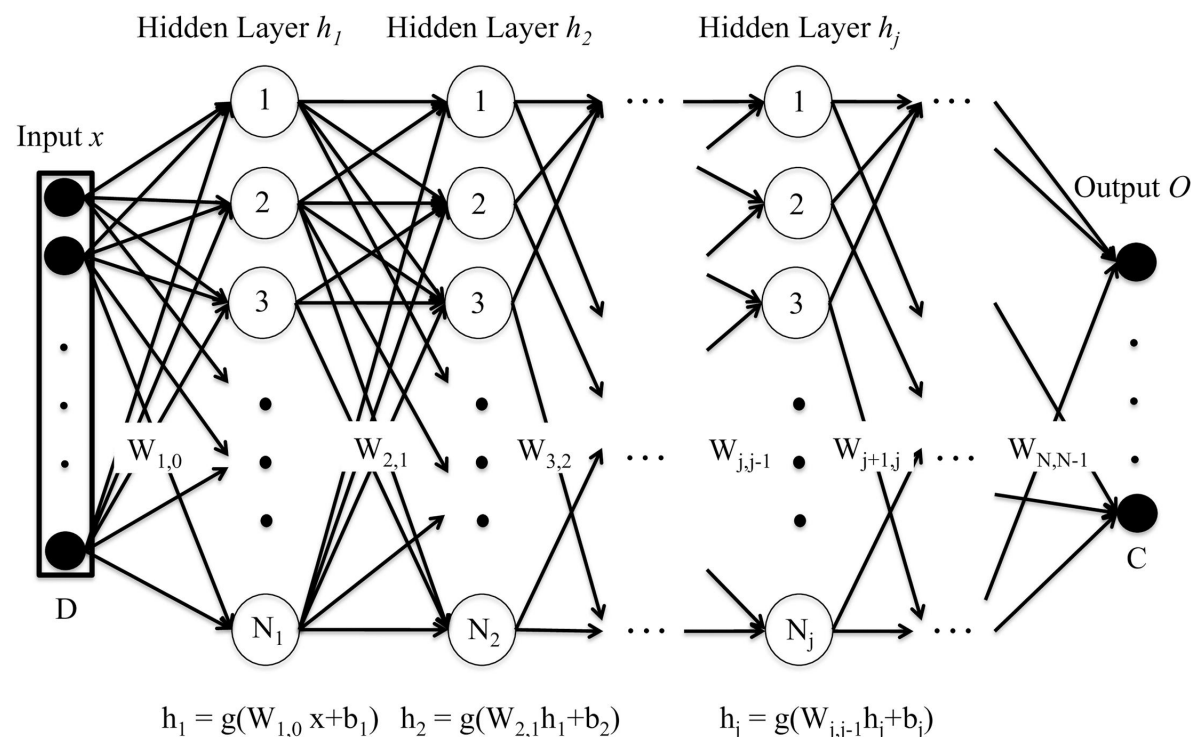
$$a^{[3]} = g^{[3]}(z^{[3]})$$

Lớp 4:

$$z^{[4]} = W^{[4]} * a^{[3]} + b^{[4]}$$

$$a^{[4]} = g^{[4]}(z^{[4]})$$

Mạng L lớp ẩn



Tổng quát, với lớp thứ l trong mạng neural ($l = 1..L$, với L là số lượng layer trong mạng neural), ta có các hàm truyền như sau:

$$z^{[l]} = W^{[l]} * a^{[l-1]} + b^{[l]}$$

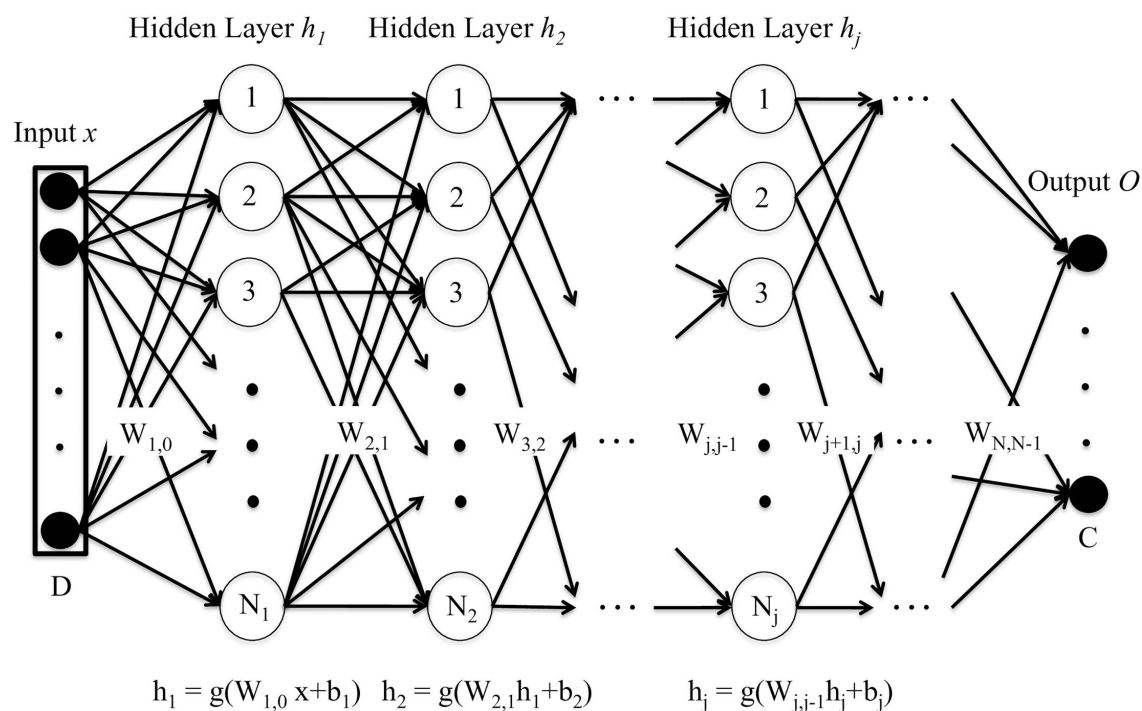
$$a^{[l]} = g^{[l]}(z^{[l]})$$

Truyền xuôi với 1 điểm dữ liệu

For l from 1 to L :

$$z^{[l]} = W^{[l]} * a^{[l-1]} + b^{[l]}$$

$$a^{[l]} = g^{[l]}(z^{[l]})$$



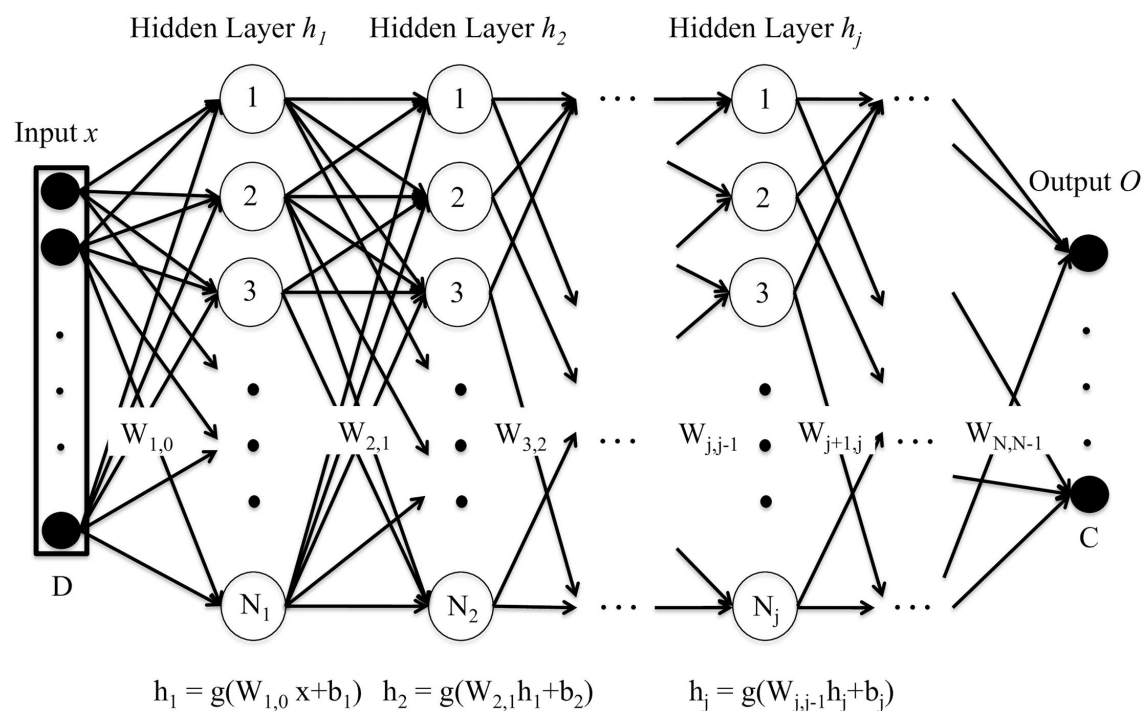
Truyền xuôi với m điểm dữ liệu

For i from 1 to m :

For l from 1 to L :

$$z^{[l]}(i) = W^{[l]}(i) * a^{[l-1]}(i) + b^{[l]}$$

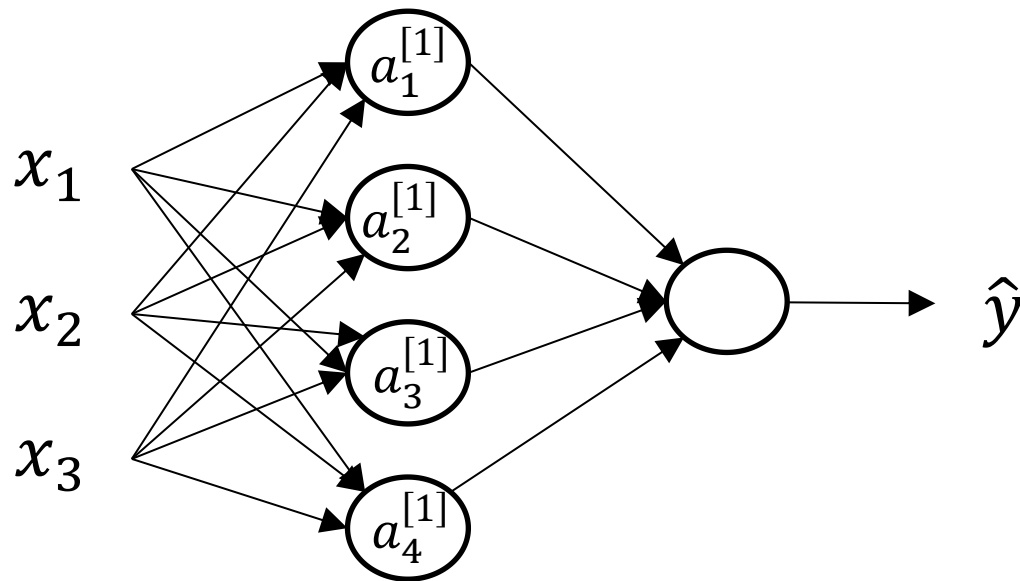
$$a^{[l]}(i) = g^{[l]}(i)(z^{[l]}(i))$$



CHIỀU CỦA MA TRẬN TRONG BIỂU DIỄN MẠNG NEURAL

MẠNG 1 LỚP ẨN – 1 điểm dữ liệu

Given input x :



Truyền xuôi
(feed forward)

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

(4x1) (4x3) (3x1) (4x1)

$$a^{[1]} = \sigma(z^{[1]})$$

(4x1) (4x1)

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

(1x1) (1x4) (4x1) (1x1)

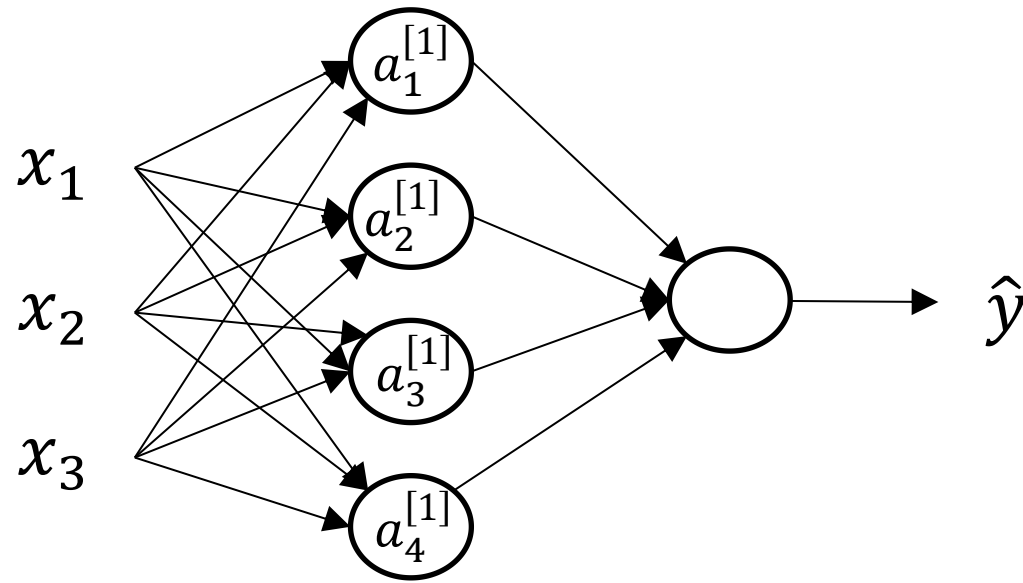
$$a^{[2]} = \sigma(z^{[2]})$$

(1x1) (1x1)

MẠNG 1 LỚP ẨN – m điểm dữ liệu

Given m input x:

For $i = 1$ to m :



$$z^{[1]}(i) = W^{[1]}(i) x^{(i)} + b^{[1]}$$

$(4 \times m) \quad (4 \times 3) \quad (3 \times m) \quad (4 \times 1)$

$$a^{[1]}(i) = \sigma(z^{[1]}(i))$$

$(4 \times m) \quad (4 \times m)$

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

$(1 \times m) \quad (1 \times 4) \quad (4 \times m) \quad (1 \times 1)$

$$a^{[2]}(i) = \sigma(z^{[2]}(i)) \rightarrow \hat{y}^{(i)}$$

$(1 \times m) \quad (1 \times m)$

MẠNG L LỚP ẨN – 1 điểm dữ liệu

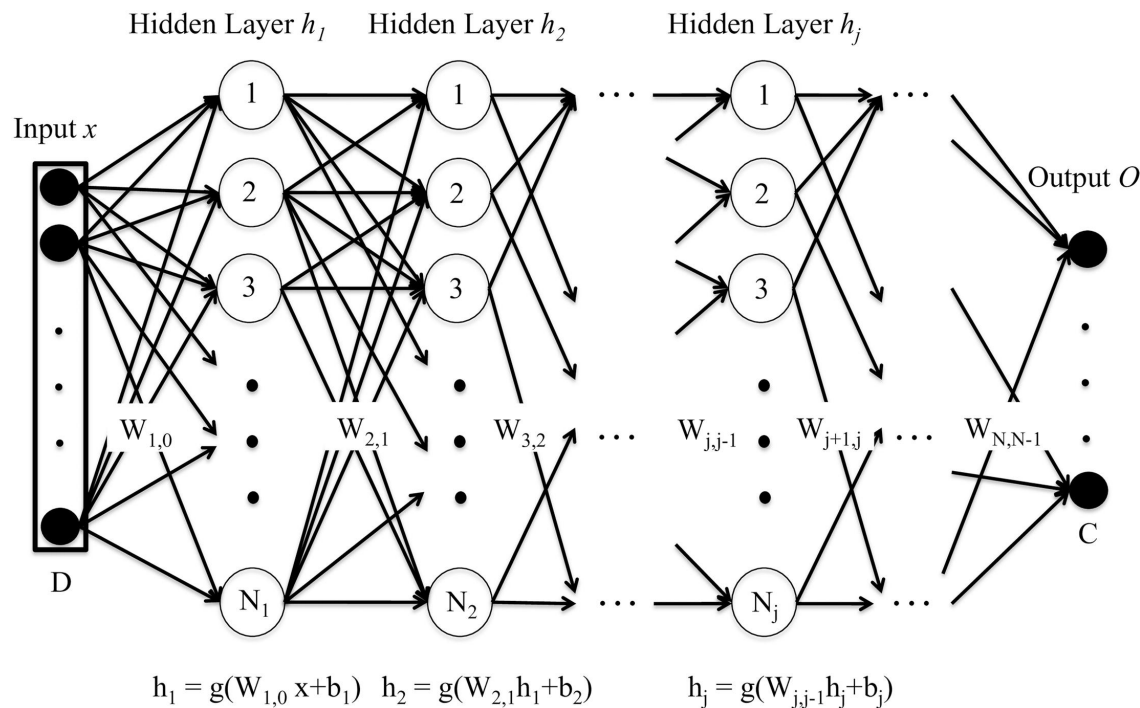
For l from 1 to L :

$$z^{[l]} = W^{[l]} * a^{[l-1]} + b^{[l]}$$

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

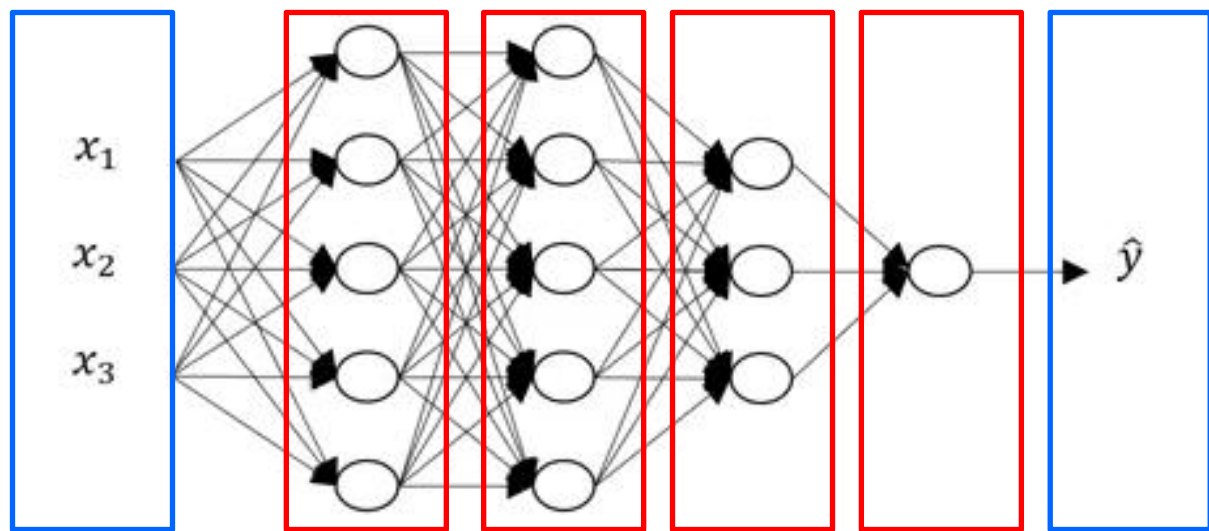
$$a^{[l]} = g^{[l]}(z^{[l]})$$

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



Ví dụ

$X = a^{[0]}$
 $\begin{matrix} z^{[1]} \\ a^{[1]} \end{matrix}$
 $\begin{matrix} z^{[2]} \\ a^{[2]} \end{matrix}$
 $\begin{matrix} z^{[3]} \\ a^{[3]} \end{matrix}$
 $\begin{matrix} z^{[4]} \\ a^{[4]} \end{matrix}$
 $\hat{y} = g^{[4]}(z^{[4]}) = a^{[4]}$



Lớp 1:

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

(4×1) (4×3) (3×1) (4×1)

$$a^{[1]} = g^{[1]}(z^{[1]})$$

(4×1) (4×1)

Lớp 2:

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

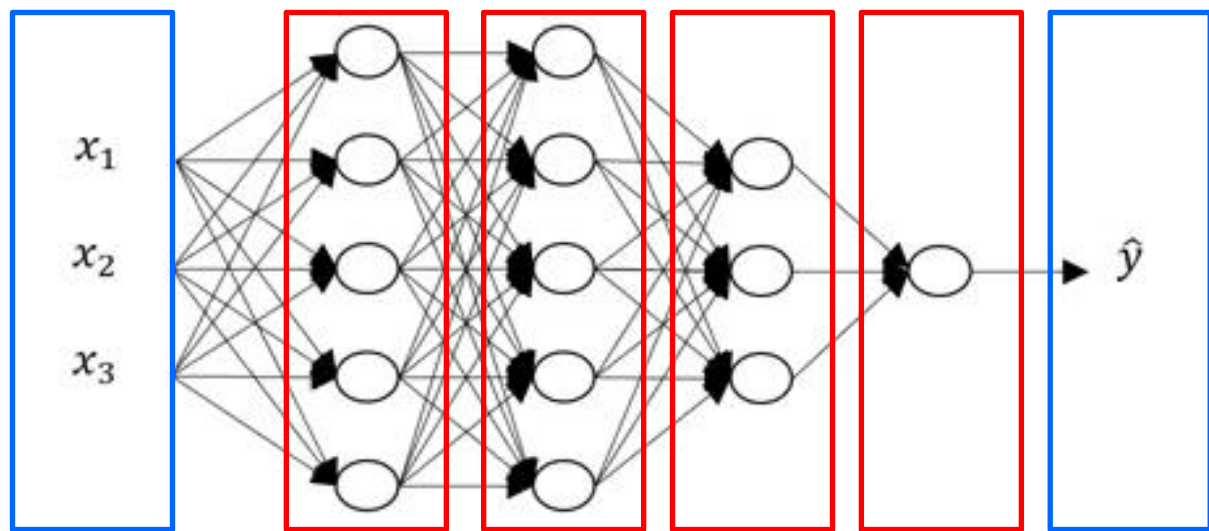
(4×1) (4×4) (4×1) (4×1)

$$a^{[2]} = g^{[2]}(z^{[2]})$$

(4×1) (4×1)

Ví dụ

$X = a^{[0]}$
 $\begin{matrix} z^{[1]} \\ a^{[1]} \end{matrix}$
 $\begin{matrix} z^{[2]} \\ a^{[2]} \end{matrix}$
 $\begin{matrix} z^{[3]} \\ a^{[3]} \end{matrix}$
 $\begin{matrix} z^{[4]} \\ a^{[4]} \end{matrix}$
 $\hat{y} = g^{[4]}(z^{[4]}) = a^{[4]}$



Lớp 3:

$$z^{[3]} = W^{[3]} * a^{[2]} + b^{[3]}$$

$(3 \times 1) \quad (3 \times 4) \quad (4 \times 1) \quad (3 \times 1)$

$$a^{[3]} = g^{[3]}(z^{[3]})$$

$(3 \times 1) \quad (3 \times 1)$

Lớp 4:

$$z^{[4]} = W^{[4]} * a^{[3]} + b^{[4]}$$

$(1 \times 1) \quad (1 \times 3) \quad (3 \times 1) \quad (1 \times 1)$

$$a^{[4]} = g^{[4]}(z^{[4]})$$

$(1 \times 1) \quad (1 \times 1)$

MẠNG L LỚP ẨN – m điểm dữ liệu

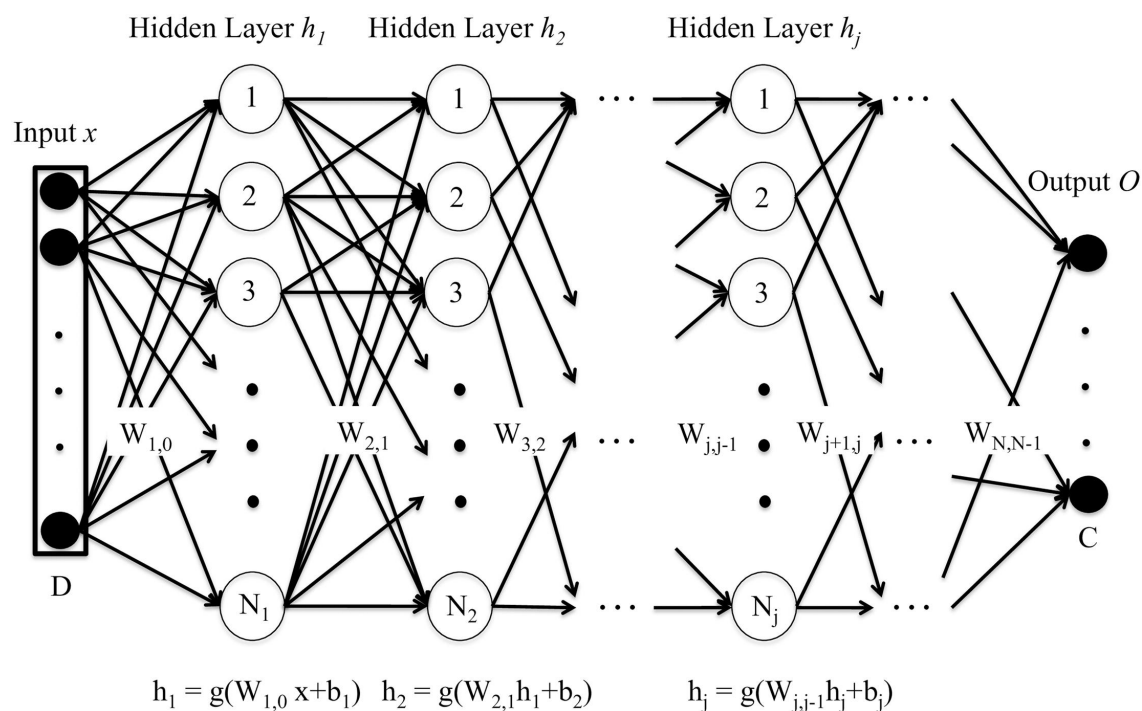
For l from 1 to L :

$$z^{[l]} = W^{[l]} * a^{[l-1]} + b^{[l]}$$

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

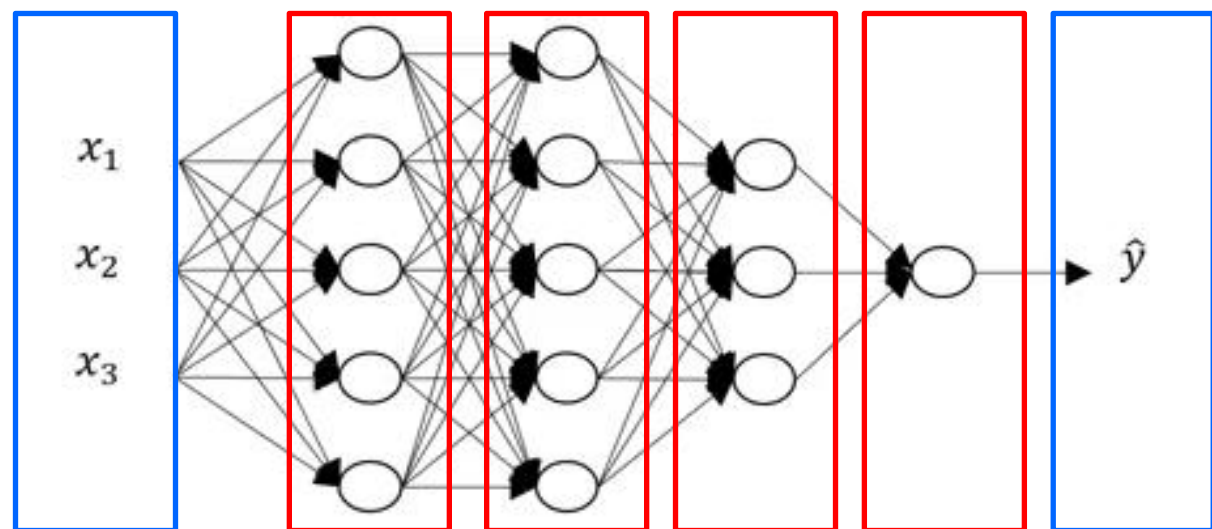
$$a^{[l]} = g^{[l]}(z^{[l]})$$

$(n^{[l]}, m)$ $(n^{[l]}, m)$



Ví dụ

$X = a^{[0]}$
 $z^{[1]}$
 $z^{[2]}$
 $z^{[3]}$
 $z^{[4]}$
 $\hat{y} = g^{[4]}(z^{[4]}) = a^{[4]}$



Lớp 1:

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

$(4 \times m)$
 (4×3)
 $(3 \times m)$
 (4×1)

$$a^{[1]} = g^{[1]}(z^{[1]})$$

$(4 \times m)$
 $(4 \times m)$

Lớp 2:

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

$(4 \times m)$
 (4×4)
 $(4 \times m)$
 (4×1)

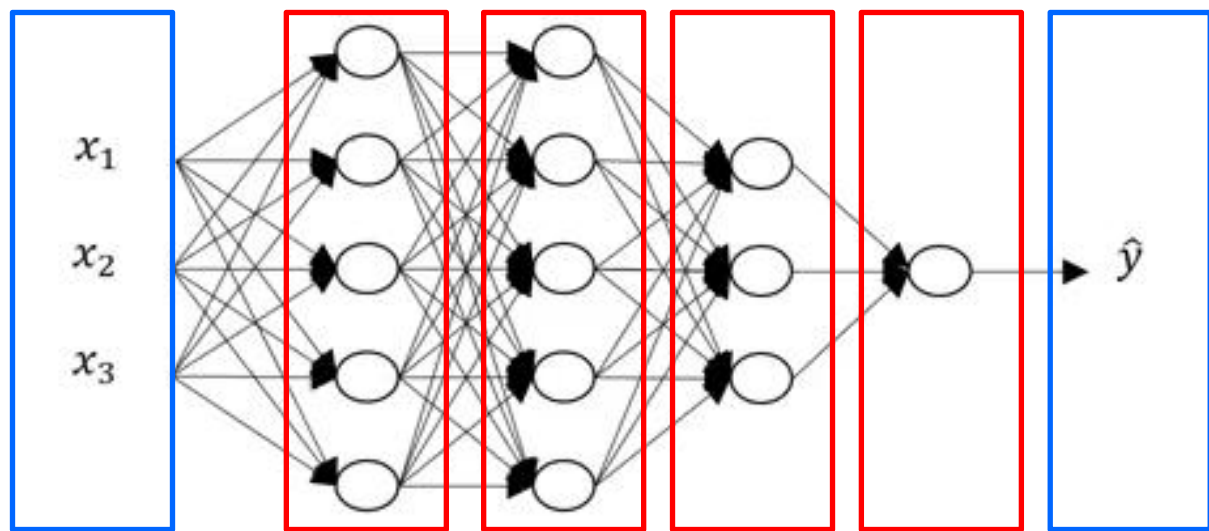
$$a^{[2]} = g^{[2]}(z^{[2]})$$

$(4 \times m)$
 $(4 \times m)$

Ví dụ

$$X = a^{[0]} \quad z^{[1]} \quad z^{[2]} \quad z^{[3]} \quad z^{[4]} \quad \hat{y} = g^{[4]}(z^{[4]}) = a^{[4]}$$

$$a^{[1]} \quad a^{[2]} \quad a^{[3]} \quad a^{[4]}$$



Lớp 3:

$$z^{[3]} = W^{[3]} * a^{[2]} + b^{[3]}$$

(3xm) (3x4) (4xm) (3x1)

$$a^{[3]} = g^{[3]}(z^{[3]})$$

(3xm) (3xm)

Lớp 4:

$$z^{[4]} = W^{[4]} * a^{[3]} + b^{[4]}$$

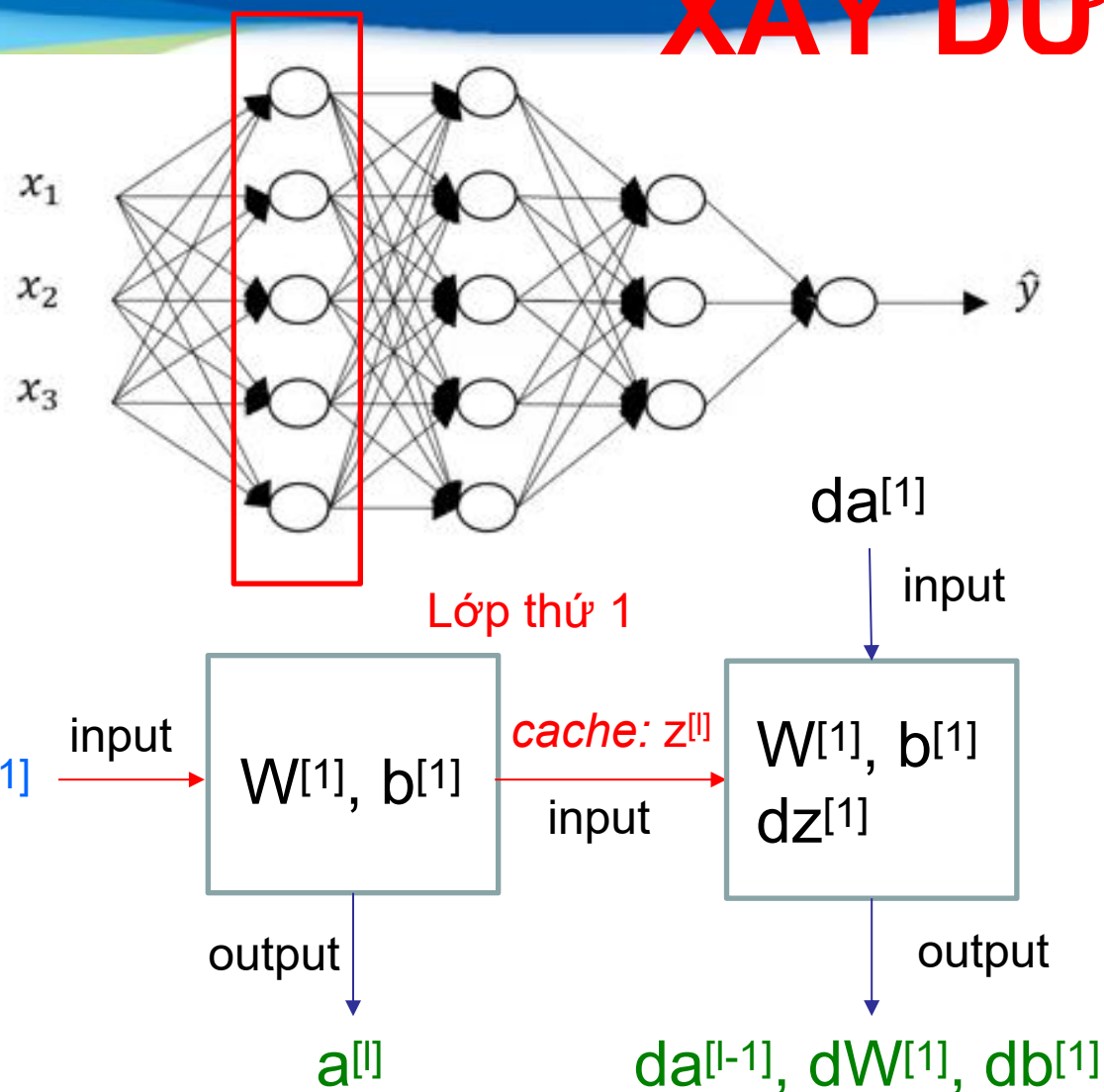
(1xm) (1x3) (3xm) (1x1)

$$a^{[4]} = g^{[4]}(z^{[4]})$$

(1xm) (1xm)

TRUYỀN NGƯỢC

XÂY DỰNG 1 LỚP



Layer thứ 1: $W^{[1]}, b^{[1]}$.

Forward:

input: $a^{[l-1]}$

output: $a^{[l]}$

cache: $z^{[l]}$

Backward:

input: $da^{[l]}$, $z^{[l]}$

output: $da^{[l-1]}, W^{[1]}, b^{[1]}$.

Forward and backward trên mạng neural nhiều lớp

Forward

- Input: $A^{[l-1]}$.
- Output: $A^{[l]}$, cached: $Z^{[l]}$.

Thực hiện:

$$Z^{[l]} = W^{[l]} * A^{[l-1]} + b^{[l]}$$

$$A^{[l]} = g^{[l]}(z^{[l]})$$

Backward

- Input: $dA^{[l]}$.
- Output: $dA^{[l-1]}$, $W^{[l]}$, $b^{[l]}$.

Thực hiện:

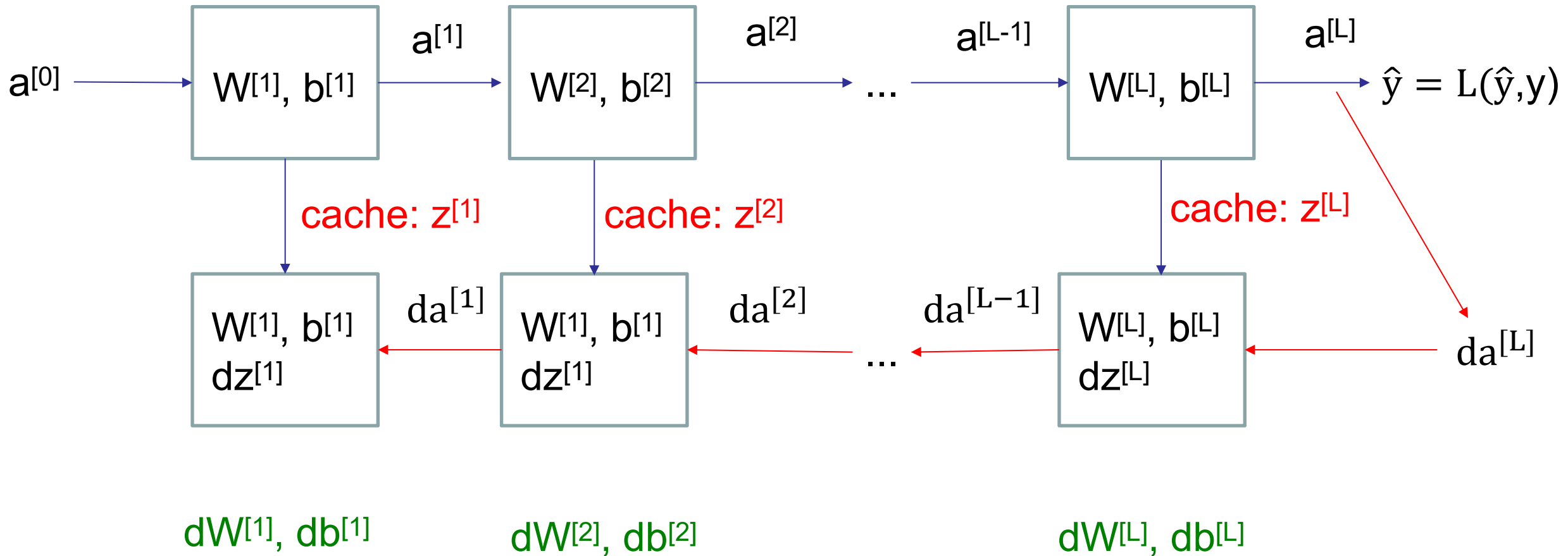
$$dZ^{[l]} = dA^{[l]} \otimes g^{[l]}(Z^{[l]})$$

$$dW^{[l]} = dZ^{[l]} * A^{[l-1]}$$

$$db^{[l]} = dZ^{[l]}$$

$$dA^{[l-1]} = W^{[l].T} * dZ^{[l]}.$$

Quá trình Forward và Backward trên nhiều lớp



Các loại tham số

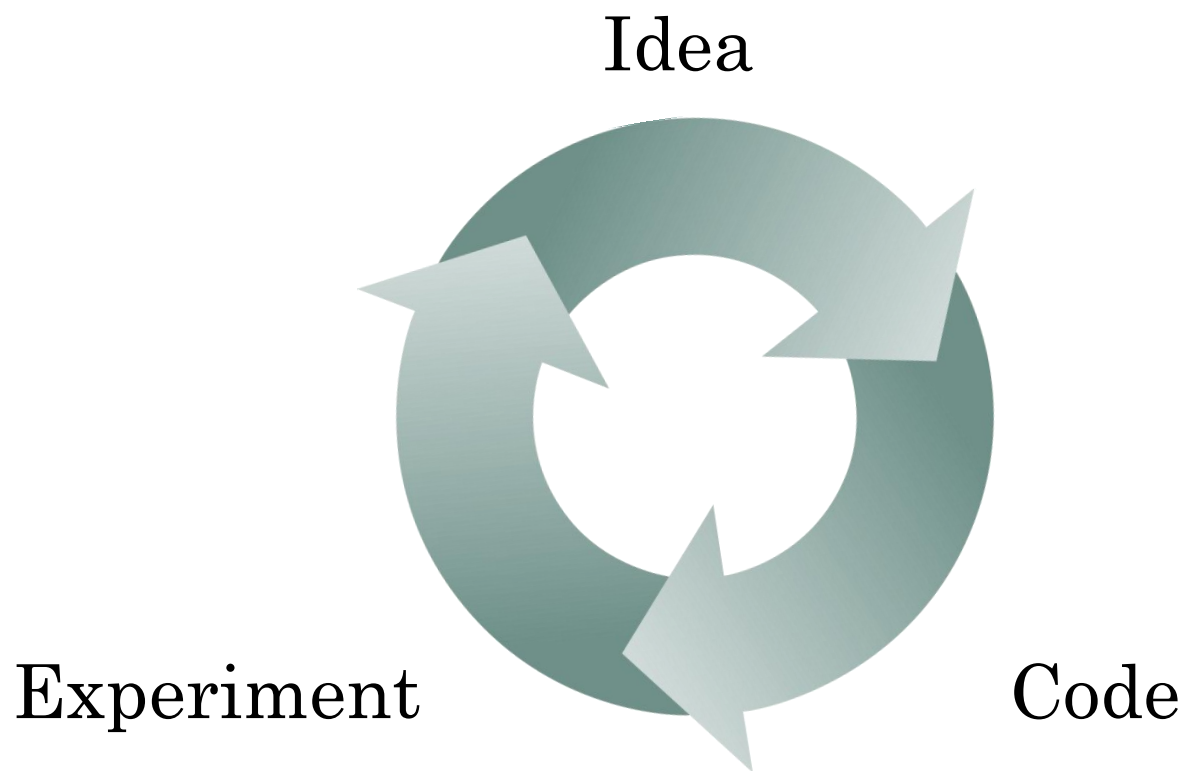
Tham số trong mô hình

- Các tham số (**parameter**) trong mô hình là: W , b .
- Các siêu tham số (**hyperparameter**) gồm có:
 - + *Tốc độ học (learning rate).*
 - + *Số lần lặp của Gradient descent (iteration).*
 - + *Số lượng lớp ẩn (hidden layer).*
 - + *Số lượng neural ẩn (hidden units).*
 - + *Hàm kích hoạt (activation function).*
- ➔ Mục đích của siêu tham số là giúp cho thuật toán học kiểm soát các giá trị tham số W và b cuối cùng.

Các siêu tham số khác trong mô hình

- Momentum.
- Epochs.
- Mini batch size.
- Regularization.
-

Điều chỉnh siêu tham số



- Quá trình áp dụng deep learning là 1 quá trình thực nghiệm.
- + Điều chỉnh các siêu tham số (ví dụ như learning rate, iteration, etc) sao cho hàm chi phí đạt giá trị nhỏ nhất.

Tại sao Deep learning lại tốt hơn so với các thuật toán truyền thống khác ?

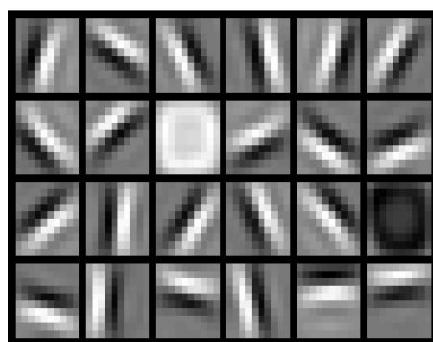
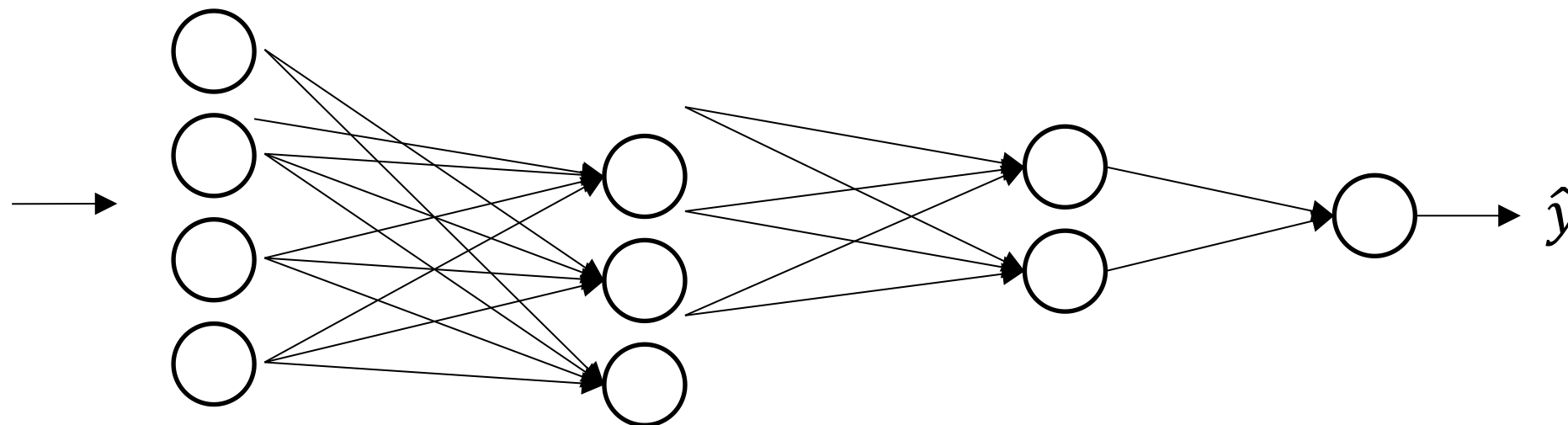
Lớp thứ 1

Lớp thứ 2

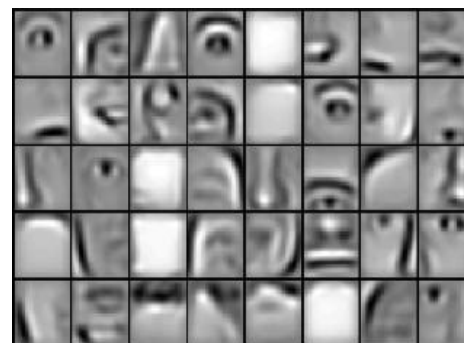
Lớp thứ 3



Dữ liệu input X



Đặc trưng 1



Đặc trưng 2



Đặc trưng 3

Kết luận

— Trong ví dụ trên:

- + **Đặc trưng thứ 1:** Chỉ là các chi tiết rất nhỏ và mờ nhạt.
 - + **Đặc trưng thứ 2:** Là các chi tiết ứng với các bộ phận trên khuôn mặt của người như: mắt, mũi, miệng, ...
 - + **Đặc trưng thứ 3:** Khuôn mặt đầy đủ, có sự phân biệt rõ rệt.
- Càng qua nhiều lớp, lượng thông tin thu được càng nhiều, càng có giá trị.
- Lớp càng sâu, lượng thông tin thu được càng có giá trị.

Tổng kết về mạng neural nhiều lớp

Backward propagation

Forward propagation

$$\begin{aligned} Z^{[1]} &= W^{[1]}X + b^{[1]} \\ A^{[1]} &= g^{[1]}(Z^{[1]}) \\ Z^{[2]} &= W^{[2]}A^{[1]} + b^{[2]} \\ A^{[2]} &= g^{[2]}(Z^{[2]}) \\ &\vdots \\ A^{[L]} &= g^{[L]}(Z^{[L]}) = \hat{Y} \end{aligned}$$

$$\begin{aligned} dZ^{[L]} &= A^{[L]} - Y \\ dW^{[L]} &= \frac{1}{m} dZ^{[L]} A^{[L]T} \\ db^{[L]} &= \frac{1}{m} np. \text{sum}(dZ^{[L]}, axis = 1, keepdims = True) \\ dZ^{[L-1]} &= dW^{[L]T} dZ^{[L]} g'^{[L]}(Z^{[L-1]}) \\ &\vdots \\ dZ^{[1]} &= dW^{[L]T} dZ^{[2]} g'^{[1]}(Z^{[1]}) \\ dW^{[1]} &= \frac{1}{m} dZ^{[1]} A^{[1]T} \\ db^{[1]} &= \frac{1}{m} np. \text{sum}(dZ^{[1]}, axis = 1, keepdims = True) \end{aligned}$$

TÀI LIỆU THAM KHẢO

1. Khoá học *Neural Network and Deep learning*, deeplearning.ai.
2. Ian Goodfellow, Yoshua Bengio, Aaron Courville, *Deep learning*, MIT Press, 2016.
3. Andrew Ng., *Machine Learning Yearning*. Link:
<https://www.deeplearning.ai/machine-learning-yearning/>
4. Vũ Hữu Tiệp, *Machine Learning cơ bản*, NXB Khoa học và Kỹ thuật, 2018.