

Assignments (by hand) : by Nguyễn Hoàng Quân

File PDF gồm 2 phần:

Phần 1 : Assignment 1: Compute adaptive weight for ANN

Phần 2 : Assignment 2: Compute adaptive weight for CNN

Assignment 1: Compute adaptive weight for ANN

Compute adaptive weight for ANN (by hand)

Given the initial weights:

| | | | | | | | |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| w _{1j} | w _{1i} | w _{2j} | w _{2i} | w _{3j} | w _{3i} | w _{jk} | w _{ik} |
| 0.2 | 0.1 | 0.3 | -0.1 | -0.1 | 0.2 | 0.1 | 0.5 |

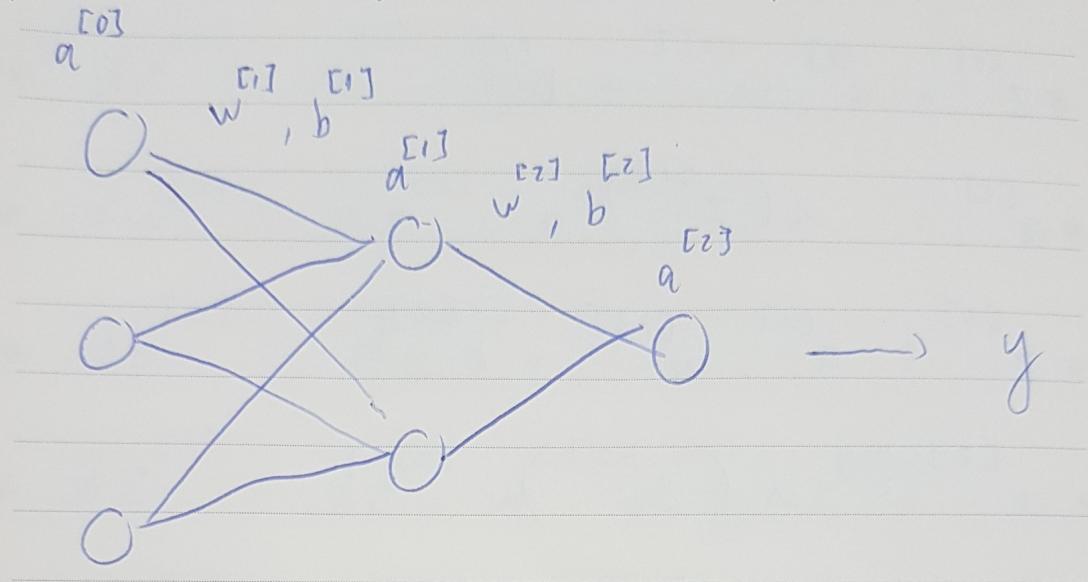
Input: [1, 0.4, 0.7]

Nodes:

(1,2,3) -> (j,i) -> k

Calculate the updated weights for the first iteration.

Computed weight for ANN



traj do:

$$* y = 0, \alpha = 0, 1 \text{ (learning rate)}$$

$$* w^{[1]} = \begin{bmatrix} 0,2 & 0,3 & -0,1 \\ 0,1 & -0,1 & 0,2 \end{bmatrix}$$

$$* w^{[2]} = \begin{bmatrix} +0,1 & 0,5 \end{bmatrix}$$

$$* b^{[1]} = \begin{bmatrix} 0 & 0 \end{bmatrix}$$

$$* b^{[2]} = \begin{bmatrix} 0 \end{bmatrix}$$

$$* a^{[0]} = \begin{bmatrix} 1 & 0,4 & 0,7 \end{bmatrix}$$

Forward prop:

$$* z^{[1]} = a^{[0]} \cdot w^{[1]T} + b^{[1]}$$

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

$$* z^{[2]} = a^{[1]} \cdot w^{[2]T} + b^{[2]}$$

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

$$* L = L(a^{[2]} - y)$$

FREE YOUR HAPPINESS

Backward prop

$$* dz^{[2]} = a^{[2]} - y$$

$$* dw^{[2]} = dz^{[2]}. a^{[1]}$$

$$* db^{[2]} = dz^{[2]}$$

$$* dz^{[1]} = dz^{[2]}. w^{[2]} * g_1'(z^{[1]})$$

$$* dw^{[1]} = dz^{[1]}. a^{[0]}$$

$$* db^{[1]} = dz^{[1]}$$

update parameter :

$$* \underline{w_1 = w_1 - d}$$

$$* w^{[2]} = w^{[2]} - \alpha \cdot dw^{[2]}$$

$$* w^{[1]} = w^{[1]} - \alpha \cdot dw^{[1]}$$

$$* b^{[2]} = b^{[2]} - \alpha \cdot db^{[2]}$$

$$* b^{[1]} = b^{[1]} - \alpha \cdot db^{[1]}$$

Áp dụng công thức trên, ta có:

Forward:

$$* Z^{[1]} = a^{[0]} \cdot w^{[1]T} + b^{[1]} = [0,25 \ 0,2]$$

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

* Vì $Z^{[1]}$ đều mang giá trị dương, ta sẽ không dùng ReLU vì nếu dùng, bài toán có thể mang sẽ bị mất đi neural network sẽ bị "tụt" tính:

\Rightarrow dùng Sigmoid

$$\Rightarrow a^{[1]} = \frac{1}{1 + \exp(-Z^{[1]})} = [0,562 \ 0,55]$$

$$* Z^{[2]} = a^{[1]} \cdot w^{[2]T} + b^{[2]} = [0,336]$$

$$* a^{[2]} = \text{Sigmoid}(Z^{[2]}) = 0,583$$

(Sigmoid với bài toán Logistic)

backward prop:

$$* dz^{[2]} = a^{[2]} \cdot y = 0,583$$

$$* dw^{[2]} = dz^{[2]} \cdot a^{[1]} = [0,327 \quad 0,32]$$

$$* db^{[2]} = dz^{[2]} = 0,583$$

$$* dz^{[1]} = dz^{[2]} \cdot w^{[2]} * \cancel{d} \underbrace{\text{sigmoid}'(z^{[1]})}_{a^{[1]}(1 - a^{[1]})}$$

$$\Rightarrow dz^{[1]} = [0,014 \quad 0,07]$$

$$* dw^{[1]} = dz^{[1]T} \cdot a^{[0]} = \begin{bmatrix} 0,014 & 0,0056 & 0,0098 \\ 0,07 & 0,028 & 0,049 \end{bmatrix}$$

$$* db^{[1]} = dz^{[1]} = [0,014 \quad 0,07]$$

Update parameter:

$$* w^{[2]} = w^{[2]} - \alpha \cdot dw^{[2]} = [0, 0.673 \quad 0, 468]$$

$$* w^{[1]} = w^{[1]} - \alpha \cdot dw^{[1]} = \begin{bmatrix} 0,1986 & 0,29944 & -0,10098 \\ 0,093 & -0,1028 & 0,1951 \end{bmatrix}$$

$$* b^{[2]} = b^{[2]} - \alpha \cdot db^{[2]} = [-0,0583]$$

$$* b^{[1]} = b^{[1]} - \alpha \cdot db^{[1]} = [-0,0014 \quad -0,007]$$

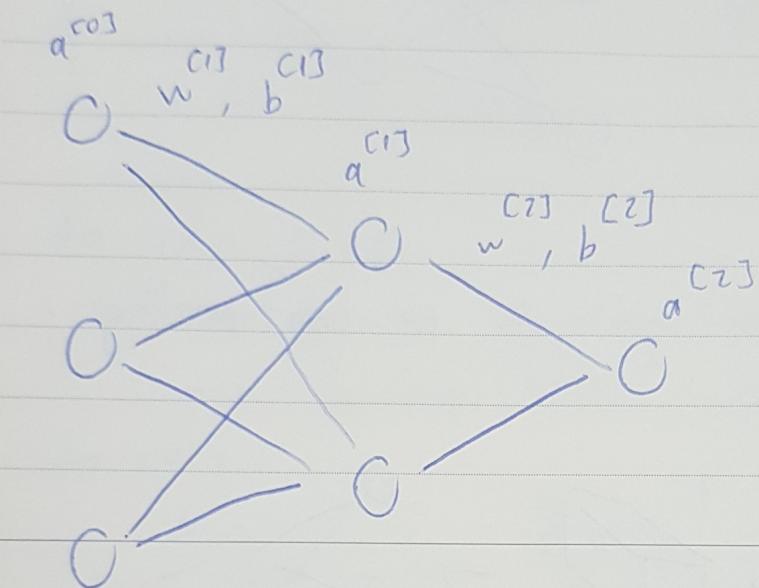
Assignment 2: Compute adaptive weight for CNN

Compute adaptive weight for CNN (by hand)

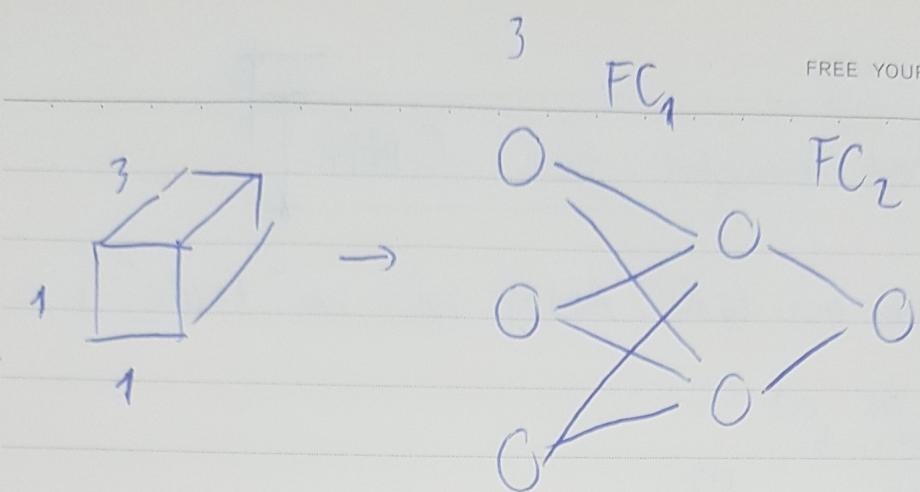
Modify the previous exercise (ANN) for a CNN which contains a convolution layer, a max-pooling layer and a fully-connected output layer.

Computed Weight for CNN

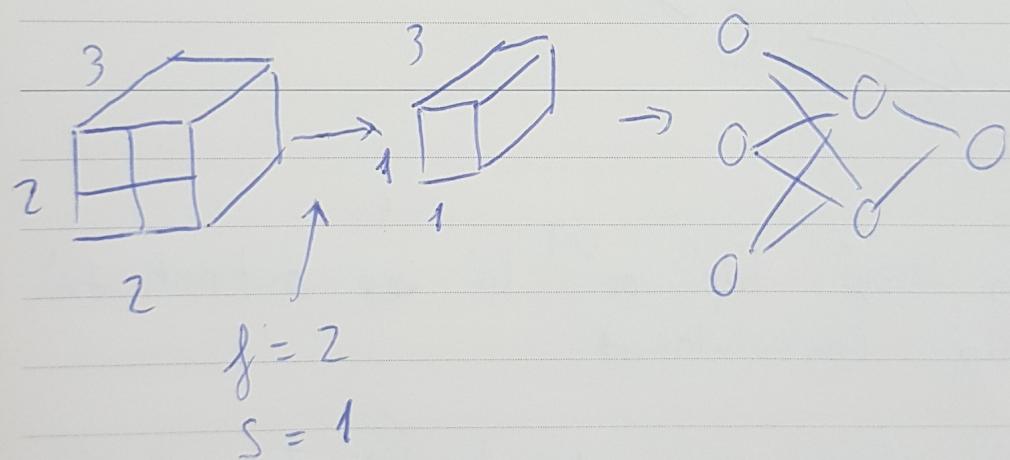
đưa vào mô hình ANN từ bài trước:



- ta thấy lớp $a^{[0]}$ là input output từ lớp Convolutional
- tuy nhiên ss' mode lai là 3, tacó theo cho lớp sau CNN trước $a^{[0]}$ như sau:



* Vì thế bài yêu cầu có lớp pooling:
ta cho rằng:



\Rightarrow output của lớp neural convolutional
là ma trận $2 \times 2 \times 3$

Màu này, ta có công thức sau:

$$\left[\frac{m + 2p - f}{s} + 1 \right] = 2$$

Ta cho $p=0$, $s=1$, $f=3$

Vì output có 3 channel

\Rightarrow 3 lớp filter

Ta có:

$$\frac{m + 2.0 - 3}{1} + 1 = 2$$

$$\Rightarrow m = 4$$

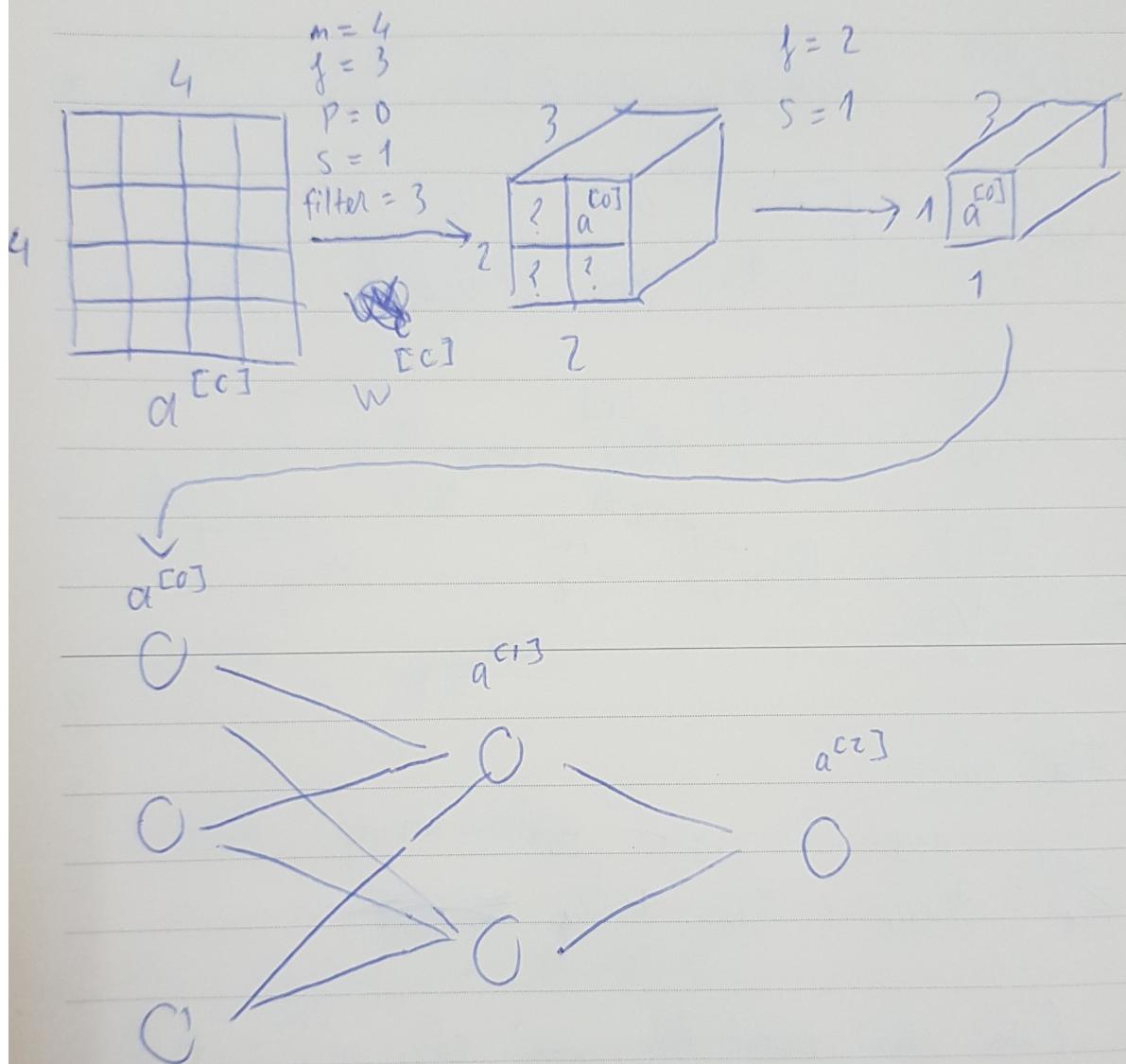
vậy ta có ma trận ban đầu 4×4

& ta cho lái chỉ có 1 channel
(trắng đen)

$$\Rightarrow \text{ma trận } a^{[0]} = 4 \times 4 \times 1$$

$$w^{[0]} = 3 \times 3 \times 3$$

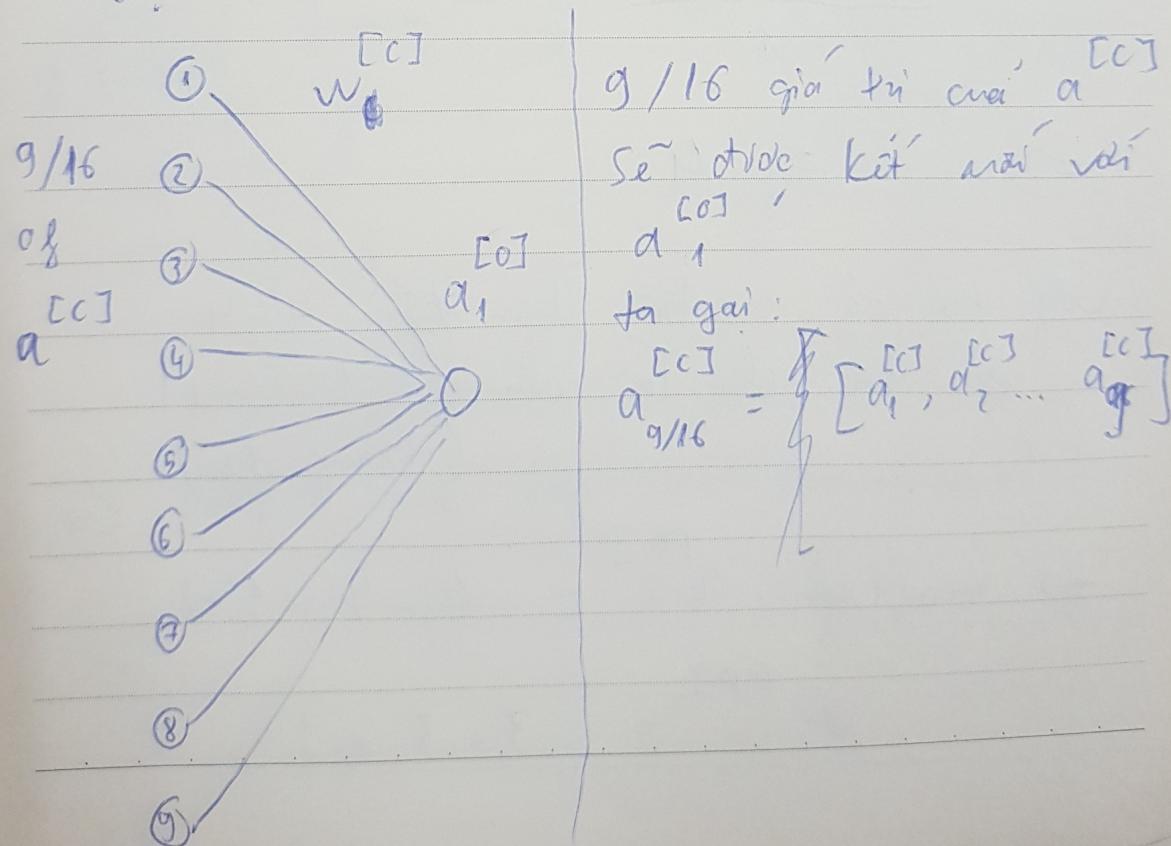
Ta có mô hình:



màu vào khía trục, ta thấy a [0] chia
lại số lần nhau để có thể được gộp
lại qua lớp Max pooling

\Rightarrow chỉ có $1/4$ của giá trị output có
đao hàn, nglwan lại update
cụ thể ở đây là 3 giá trị cui a [0]

\Rightarrow trong bài này, ta chỉ làm ví dụ 1 giá
trị, 2 giá trị còn lại tự đồng tự:
 \Rightarrow giá trị a [0]



Ta có:

$$* a^{[0]} = [1 \ 0,4 \ 0,7]$$

$$* d^{[1]} = [0,014 \ 0,07]$$

$$* w^{[1]} = \begin{bmatrix} 0,2 & 0,3 & -0,1 \\ 0,1 & -0,1 & 0,2 \end{bmatrix}$$

$$* a_{g/16}^{[c]} = [a_1^{[c]} \ a_2^{[c]} \ \dots \ a_g^{[c]}]$$

$$* w_{1/3}^{[c]} = [w_1^{[c]} \ w_2^{[c]} \ \dots \ w_g^{[c]}]$$

Ta có:

$$z_1^{[0]} = a_{g/16}^{[c]} \cdot w_{1/3}^{[c] T} + b_1 \quad \cancel{\text{tuy}} \cancel{\text{m}}$$

(ta cho rằng hằng số $a_1^{[c]} = 1$)

$$\Rightarrow a_1^{[c]} = 1 = \text{Sigmoid}(z_1^{[0]})$$

$$\Rightarrow z_1^{[0]} = \infty$$

⇒ output của $a^{[c]}$ là ∞ (điều này là sai)

⇒ chọn ReLU

$$\Rightarrow z_1^{[0]} = 1$$

Back prop :

$$* dz^{[0]} = dz^{[1]}. w^{[1]} * \underbrace{g_o'(z^{[0]})}_{=1}$$

$$\Rightarrow dz^{[0]} = [0,0098 \quad -0,0028 \quad 0,0126]$$

$$* dw_{1/3}^{[c]} = dz^{[0]T} \cdot a_{g/16}^{[c]}$$

$$* db_{1/3}^{[c]} = dz^{[0]}$$

Update parameters:

$$* w_{1/3}^{[c]} = w_{1/3}^{[c]} - \alpha \cdot dw_{1/3}^{[c]}$$

$$* b_{1/3}^{[c]} = b_{1/3}^{[c]} - \alpha \cdot db_{1/3}^{[c]}$$

\Rightarrow Khi có giá trị khai đầu của $a^{[c]}$
 và giá trị của $w^{[c]}$, $b^{[c]}$
 Ta thế vào công thức để ra kết quả