Backpropagation

(Numerical Ex)

Binary — دو کلاسه ⟶ سگ هست / سگ نیست

احتمال سگ بودن → [1]

Sigmoid → [ ] ← احتمال سگ بودن

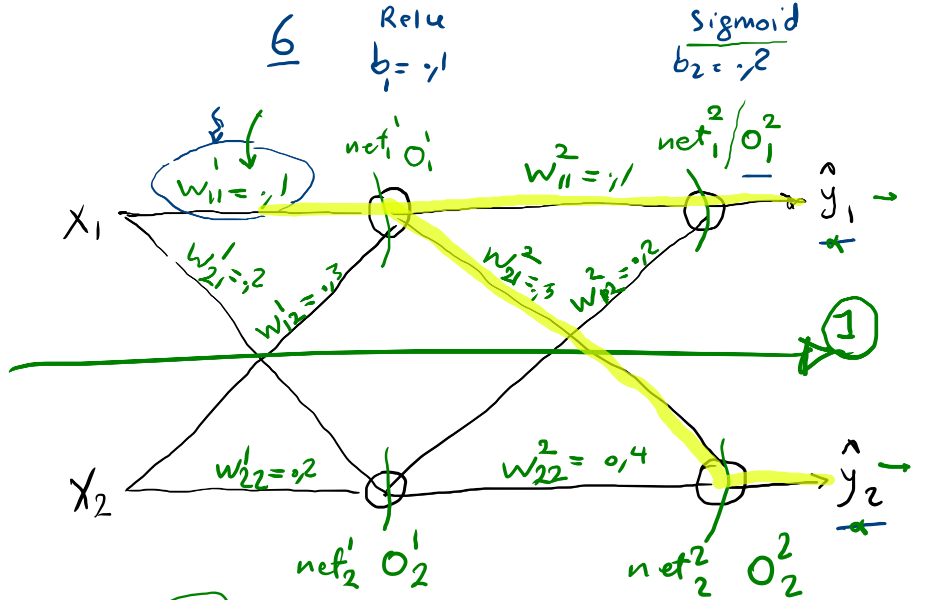
Binary — تصویر ⟶ احتمال گربه بودن ○ / احتمال سگ بودن ○ → Dense(2, activation = 'softmax')

$$\begin{bmatrix} .3 \\ .7 \end{bmatrix}$$

مثل سگ و گربه

Sigmoid با → □ سگ → 1 / □ گربه → 0 ⟶ Train — NN ⟶ output (Dens(1, active='sigmoid')

Neuron



X_1	X_2
0,7	0,5

$$\begin{cases} y_1\,\text{True} = 0{,}3 \\ y_2\,\text{True} = 0{,}7 \end{cases}$$

1.

$$net_1^1 = w_{11}^1 X_1 + w_{12}^1 X_2 + b_1$$

$$O_1^1 = \text{Relu}(net_1^1)$$

$$net_2^1 = w_{21}^1 X_1 + w_{22}^1 X_2 + b_1$$

$$O_2^1 = \text{Relu}(net_2^1)$$

$$net_1^2 = w_{11}^2 O_1^1 + w_{12}^2 O_2^1 + b_2$$

$$O_1^2 = \sigma(net_1^2) \longrightarrow \hat{y}_1$$

$$net_2^2 = w_{21}^2 O_1^1 + w_{22}^2 O_2^1 + b_2$$

$$O_2^2 = \sigma(net_2^2) \longrightarrow \hat{y}_2$$

② $L = MSE = \frac{1}{N} \sum_{i=1}^{N} (y_{i_{true}} - \hat{y}_i)^2$

SGD

به ازای هر sample که داریم
شبکه‌ی عصبی باید، یکبار وزن‌هاش
آپدیت بشود!

$\xrightarrow{SGD}$

$$L = (y_{true} - \hat{y})^2$$

○ $\hat{y}_1$

○ $\hat{y}_2$

$\left\{ \begin{matrix} \hat{y}_1 \\ \hat{y}_2 \end{matrix} \right.$

$$L = \frac{1}{2} \left[(0{,}3 - \hat{y}_1)^2 + (0{,}7 - \hat{y}_2)^2 \right]$$

$$\frac{\partial L}{\partial \hat{y}_2} = \frac{-1}{\cancel{2}} (\cancel{2}) (0{,}7 - \hat{y}_2)$$

$$\frac{\partial L}{\partial \hat{y}_1} = \frac{-1}{\cancel{2}} (\cancel{2}) (0{,}3 - \hat{y}_1)$$

③ Backpropagation

target : w_{11}^1

$$w_{11\text{-new}}^1 = w_{11\text{-old}}^1 - \alpha \frac{\partial L}{\partial w_{11}^1}$$

0/01

?

$$\frac{\partial L}{\partial w_{11}^1} = \begin{cases} \dfrac{\partial L}{\partial \hat{y}_1}\dfrac{\partial \hat{y}_1}{\partial o_1^2}\dfrac{\partial o_1^2}{\partial net_1^2}\dfrac{\partial net_1^2}{\partial o_1^1}\left[\dfrac{\partial o_1^1}{\partial net_1^1}\dfrac{\partial net_1^1}{\partial w_{11}^1}\right] \\ \dfrac{\partial L}{\partial \hat{y}_2}\dfrac{\partial \hat{y}_2}{\partial o_2^2}\dfrac{\partial o_2^2}{\partial net_2^2}\dfrac{\partial net_2^2}{\partial o_1^1}\left[\dfrac{\partial o_1^1}{\partial net_1^1}\dfrac{\partial net_1^1}{\partial w_{11}^1}\right] \end{cases}$$

?

$$\frac{\partial Relu(net_1^1)}{\partial net_1^1}$$

X_1

$$\frac{\partial L}{\partial w_{11}^1} = \left[\frac{\partial L}{\partial \hat{y}_1}\frac{\partial \hat{y}_1}{\partial o_1^2}\frac{\partial o_1^2}{\partial net_1^2}\frac{\partial net_1^2}{\partial o_1^1} + \frac{\partial L}{\partial \hat{y}_2}\frac{\partial \hat{y}_2}{\partial o_2^2}\frac{\partial o_2^2}{\partial net_2^2}\frac{\partial net_2^2}{\partial o_1^1}\right]\left[\frac{\partial o_1^1}{\partial net_1^1}\frac{\partial net_1^1}{\partial w_{11}^1}\right]$$

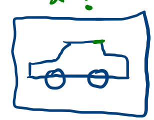
$(\hat{y}_1 - y_1)$ $\quad 1 \quad$ $\dfrac{\partial 6(net_1^2)}{\partial net_1^2}$? $\quad w_{11}^2 \quad$ $(\hat{y}_2 - y_2)$ $\quad 1 \quad$ $\dfrac{\partial 6(net_2^2)}{\partial net_2^2}$? $\quad w_{21}^2$

MSE $\rightsquigarrow$ loss function $\rightsquigarrow$ regression!

cross-entropy $\rightsquigarrow$ classification!

0 → cad! احتمال
0 → lam احتمال

x ?



NN

output

$\begin{bmatrix} 0.3 \\ 0.7 \end{bmatrix}$

1. cadillac
2. lamborghini

True : $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ → cadillac, → lamborg—!

$$\rightsquigarrow \ell = -\sum_{i}^{N} y_i \cdot \log \hat{y}_{i\,pred}$$

N: Number of categ.

$$\text{loss} = -[y_1 \log y_{1\,pred} + y_2 \log y_{2\,pred}]$$

$$\text{loss} = -[0 \cancel{\log 0.3} + 1 \log(0.7)] = -\log(0.7) =$$

$$-(-0.15) = 0.15$$

NN

cadullac

lamborgini

$$\begin{bmatrix} .4 \\ .6 \end{bmatrix} \quad \begin{bmatrix} .1 \\ .9 \end{bmatrix} \quad \begin{array}{c} \text{true} \\ \begin{bmatrix} 0 \\ 1 \end{bmatrix} \end{array}$$

1 2

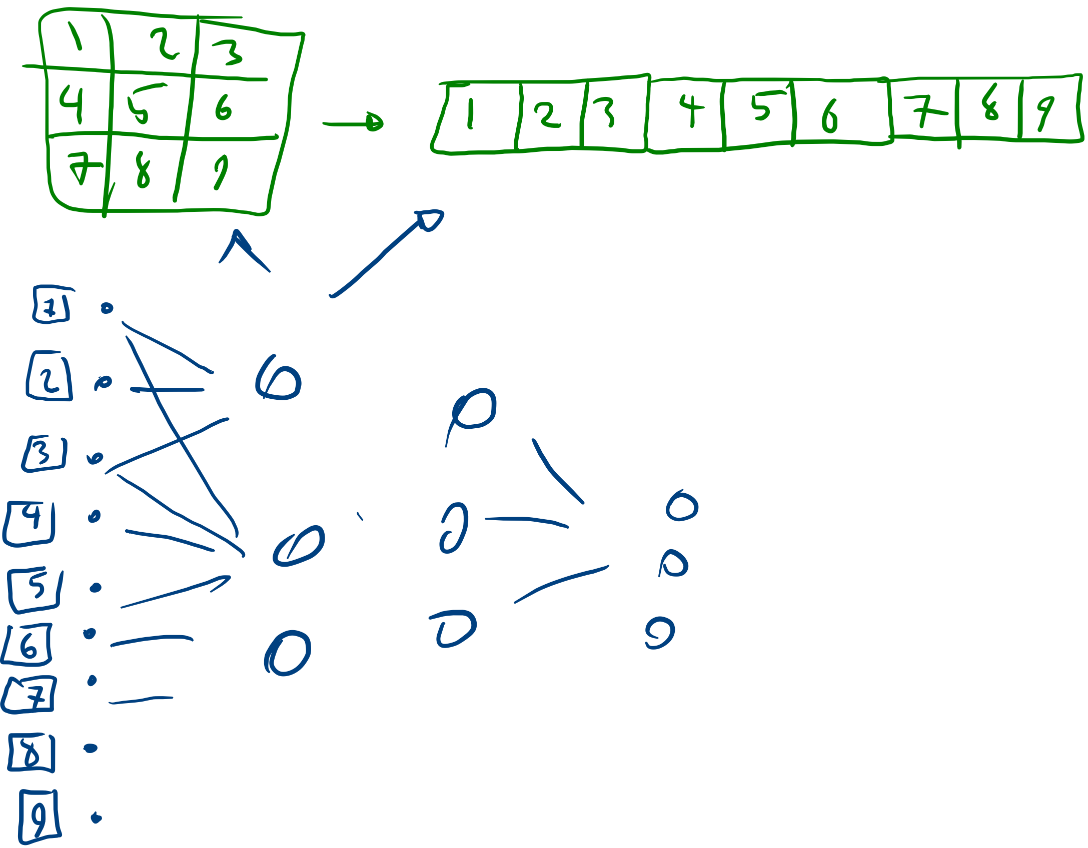
$$\text{cross-entropy-1} = -[1 \log .6] = -(-0,22) = 0,22$$

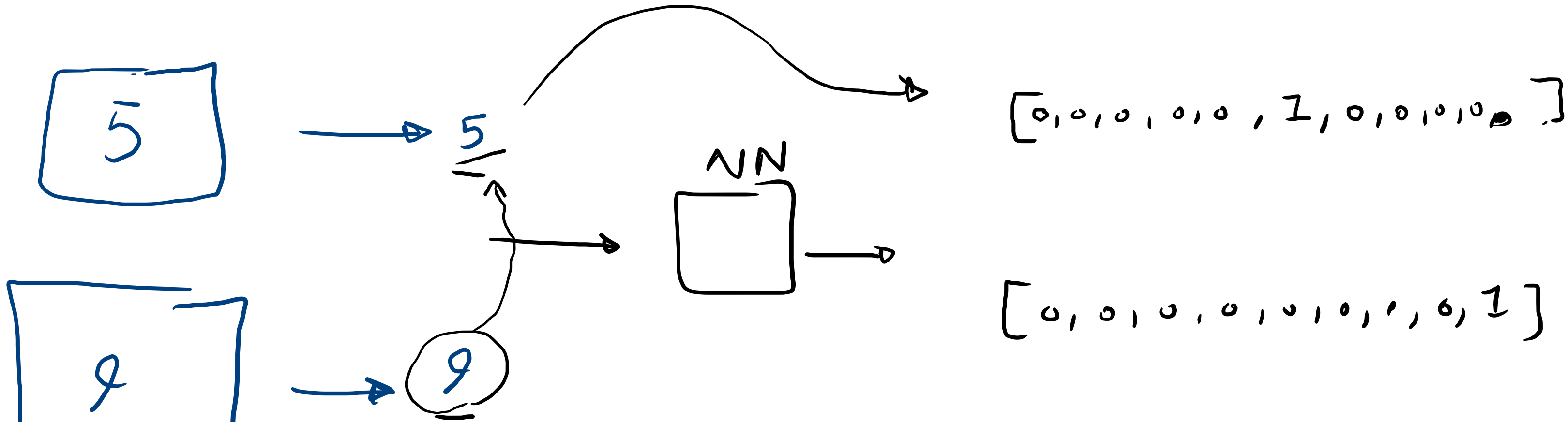
$$\text{cross-entropy-2} = -[1 \log .9] = -(-0,045) = 0,045$$

Number of category ← N

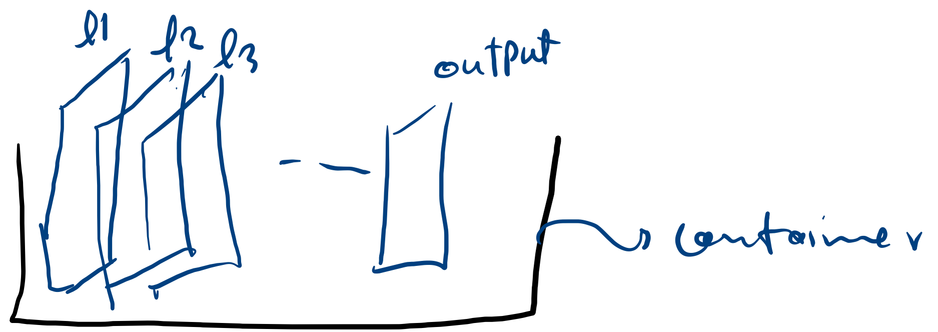
$$\text{cross} = -\sum_{i=1}^{N} y_{i\,true} \log y_{i\,pred}$$

Dense → Image !





Sequential →



~~Relu~~ $\left\{ \frac{\partial x}{\partial x} = 1 \right.$

$$\frac{\partial Relu(x)}{\partial x} = ?$$

$$Relu(x) = \begin{cases} x & x \geq 0 \\ 0 & x < 0 \end{cases}$$

$$\frac{\partial Relu(x)}{\partial x} = \begin{cases} 1 & x > 0 \\ 0 & x < 0 \end{cases}$$

$$\boxed{x = 0 \rightarrow \text{undefined}}$$

$$\frac{\partial Relu(net_1^1)}{\partial net_1^1} = \begin{cases} 1 & net_1^1 > 0 \\ 0 & net_1^1 < 0 \end{cases}$$

$$\frac{\partial \sigma(x)}{\partial x} = \sigma(x)\left(1 - \sigma(x)\right)$$

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

$$x = net_2^2 \Rightarrow \underline{\sigma(net_2^2)}\left(1 - \sigma(\underline{net}_2^2)\right)$$

$$net_2^2 = 0.2$$

$$= \sigma(0.2)\left(1 - \sigma(0.2)\right)$$

End