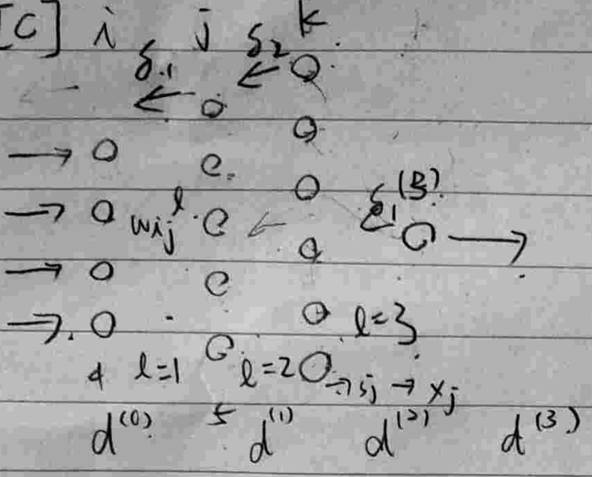
No.
Date :

ML HW6 B06611032 武敬祥

1. [C]



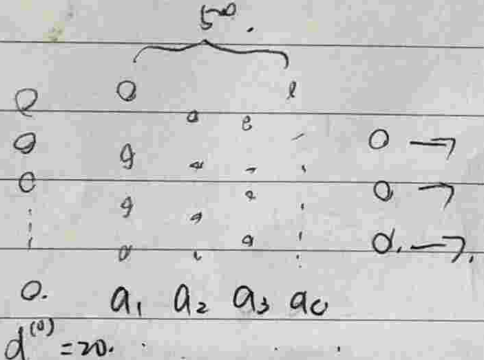
$l \in \{1, 2\}$

$$\delta_j^{(2)} = \frac{\partial e_n}{\partial s_j^{(2)}} = \underbrace{\sum_{k=1}^{k=6} (\delta_k^{(3)})(w_{jk}^{(3)})}_{5\times 6} h'(s_j^{(2)})$$

$$\delta_i^{1} = \frac{\partial e_n}{\partial s_i^{(1)}} = \underbrace{\sum_{j=1}^{j=5} (\delta_j^{(2)})(w_{ij}^{(2)})}_{5\times 4} h'(s_i^{1})$$

$5\times 6 + 5\times 4 = 50.$

2. [d]



$a_1 + a_2 + a_3 + a_4 = 50.$

find max. $(20\times a_1)+(a_1\times a_2)\cdots(a_N\times 3)$ $= \max[(20\times a_1)+(a_1\times 3)] + \max[(a_1\times a_2)+\cdots (a_{N-1})(a_N)]$ $N-1$

$$\sqrt[N-1]{(a_1\times a_2)(a_2\times a_3)\cdots(a_{N-1}a_N)} \le \frac{a_1 a_2 + a_2 a_3 \cdots a_{N-1}a_N}{N-1}$$

"=" 成立 $a_1 a_2 = a_2 a_3 = \cdots = a_{N-1}a_N$ $\Rightarrow a_1 = a_3 = a_5 \cdots = a_a$

$a_1 a_2$ $a_3 a_4$ $a_4 a_5$ $a_2 = a_4 = a_6 \cdots = a_b$

max $(20\times a_a) + (N-1)(a_a a_b) + (a_N\times 3)$

偶×奇=偶
偶×偶=偶
奇×奇=奇

偶×□=偶

20

Date

3. [d]

$$q = X^{(L)} = \left[\frac{\exp(s_1^{(L)})}{\sum_{k=1}^{K} \exp(s_1^{(L)})}, \frac{\exp(s_2^{(L)})}{\sum_{k=1}^{K} \exp(s^{(L)})} \cdots \frac{\exp(s_K^{(L)})}{\sum_{k=1}^{K} \exp(s_k^{(L)})} \right]$$

$$v = [[y=1], \cdots [y=k]]$$

$$err(x, y) = -\sum_{k=1}^{K} v_k \ln q_k$$

$$\frac{\partial err(x,y)}{\partial s_k^{(L)}} = \frac{-\sum_{k=1}^{K} v_k \ln \frac{\exp(s_k^{(L)})}{\sum_{k=1}^{K} \exp(s_k^{(L)})}}{\partial s_k^{(L)}} = \frac{-\sum_{k=1}^{K} v_k \ln \exp(s_k^{(L)})}{\partial s_k^{(L)}} + \frac{\sum_{k=1}^{K} v_k \ln \sum_{k=1}^{K} \exp(s_k^{(L)})}{\partial s_k^{(L)}}$$

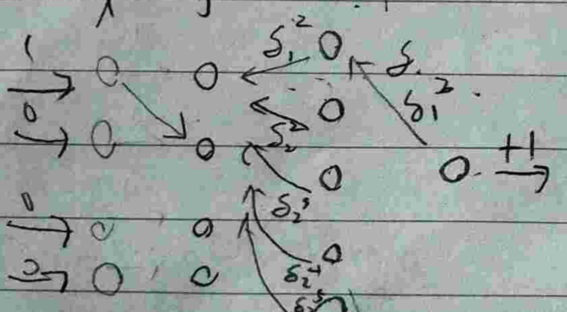
$$= \frac{-\sum_{k=1}^{K} v_k s_k^{(L)}}{\partial s_k^{(L)}} + \frac{\sum_{k=1}^{K} v_k \ln(e^{s_1^{(L)}} + e^{s_2^{(L)}} \cdots)}{\partial s_k^{(L)}} \qquad \frac{\ln e^{s_1(L)} \cdots}{e^{s_1(L)} \cdots} \qquad \frac{e^{s_1(L)} \cdots}{\partial s_k}$$

$$= -v_k + \left(e^{s_k^{(L)}} \cdot 1 \cdot \frac{1}{\sum_{k=1}^{K} \exp(s_k^{(L)})}\right) \qquad e^{s_1^{(L)}} \cdots \cdot 1 \cdot$$

$$= -v_k + q_k \ \#$$

4. [a] $l=1 \quad l=2 \quad l=3$

$i \quad j \quad k$



update rule:

$$w_{ij}^{(l)} \leftarrow w_{ij}^{(l)} - \eta x_i^{(l-1)} \cdot \delta_j^{(l)}$$

$$w_{01}^{(1)} \leftarrow w_{01}^{(1)} - \eta x_0^{0} \cdot \delta_1^{(1)}$$

$$\delta_1^{1} = \frac{\partial e_n}{\partial s_1^{1}} = \sum_{k=1}^{5} (\delta_k^{2})(w_{1k}^{2})' \tanh'(s_1^{1})$$

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

$$\tanh'(x) = \frac{4}{(e^x + e^{-x})^2}$$

$$x_n = (1, 0, 0, 0)$$

$$\frac{\partial e_n}{\partial w_{k1}} = \frac{\partial (y_n - \sum_{k=1}^{5} w_{k1} \cdot x_k^2)^2}{\partial w_{k1}}$$

$$w_{k1}^{(3)} = 0 \Rightarrow s_k^{(2)} = 0 \Rightarrow w_{1k}^{(2)} = 0$$

$$\Rightarrow s_1^{1} = 0 \Rightarrow w_{01}^{(1)} = 0$$

$$w_{1k}^{(2)} \leftarrow w_{1k}^{(2)} - \eta x_1^{1} \cdot \delta_k^{2}.$$

$$\delta_k^{(2)} = \frac{\partial e_n}{\partial s_k^{(2)}} = \delta_1^{3} (w_{k1}^{(3)}) (\tanh'(s_k^{(2)}))$$

$$w_{k1}^{(3)} = w_{k1}^{(3)} - \eta x_k^{2} \cdot \delta_1^{3}.$$

$$\delta_1^{3} = -2(y_n - s_1^{(L)}) = -2(1-0) = -2$$

Chrys culture

No.

Date

7. [d].

$G(x) = \text{sign}\left(\sum_{t=1}^{3} 1\cdot g_t(x)\right)$

$E_{out}(G) = 0.2$

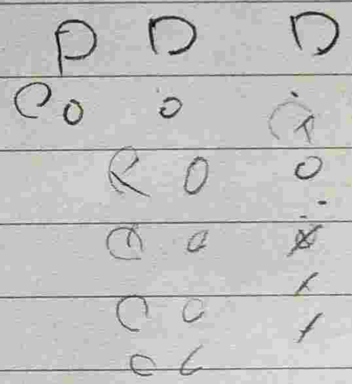
$g_1(x), g_2(x), g_3(x)$

at least two $g(x)$ raise error. for $E_{out}(G) = 0.2$

$\sum_{t=1}^{T} E_{out}(g_t(x)) \ge 0.4$

if cases is correct for G. $1 - E_{out}(G) = 0.8$.

at least two $g(x)$ is correct.

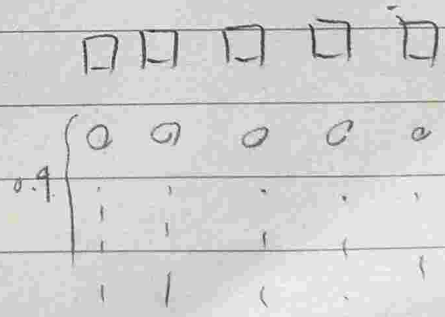


$\frac{4}{5}\times 1 + \frac{1}{5}\times 3 \ge \sum_{t=1}^{T} E_{out}(g_t(x))$

$1.4 \ge \sum_{t=1}^{T} E_{out}(g_t(x)) \ge 0.4$

8. [c]

$E_{out}(g_t) = 0.4 \quad \forall t \in \{1, 2, 3, 4, 5\}$



$G(x) = \text{sign}\left(\sum_{t=1}^{5} g_t(x)\right)$

for $X = \{x_1, x_2, \dots x_n\}$

$C_3^5 (0.4)^3 (0.6)^2 + C_4^5 (0.4)^4\, 0.6 + C_5^5 \times (0.4)^5$

$= 10 \times 0.36 \times 0.064 + 5 \times 0.6 \times 0.0256 + 0.01024$

$= 0.31744$

No.
Date

9. [b]

$x = \{x_1, x_2 \dots x_N\}$

\# of $x_s = \frac{N}{2}$

if x_i is not chosen:

$$\lim_{N\to\infty} \left(\frac{N-1}{N}\right)^{\frac{N}{2}} = \lim_{N\to\infty}\left(1-\frac{1}{N}\right)^{\frac{N}{2}} = \left(\frac{1}{e}\right)^{0.5} \approx 0.60.$$

10.

$x: \{x_1, x_2 \dots x_d\}\ \forall x_i \in$ even

$g_{s,i,\theta}(x) = s\cdot \mathrm{sign}(x_i - \theta) \rightarrow$ odd

$$\phi_{ds}(x) = \big(g_{1,1,2L+1}(x),\ g_{1,1,2L+3}(x) \dots g_{1,1,2R-1}(x)$$
$$g_{-1,1,2L+1}(x),\ g_{-1,1,2L+3}(x) \dots g_{-1,1,2R-1}(x)$$
$$\vdots$$
$$g_{1,d,2L+1}(x) \dots g_{1,d,2R-1}(x)$$
$$g_{-1,d,2L+1}(x) \dots g_{-1,d,2R-1}(x) \big)$$

$K_{ds}(x, x') =$

No.
Date : :

19. [e]

neural networks & deep learning.
because the reason why I step into ML is that I trained a model with CNN to identify the apple. NN really astonished me.

20. [b]

matrix factorization.
I think. this topic is a little abstract to me. maybe the partial reason is that I didn't pay full attention on it.