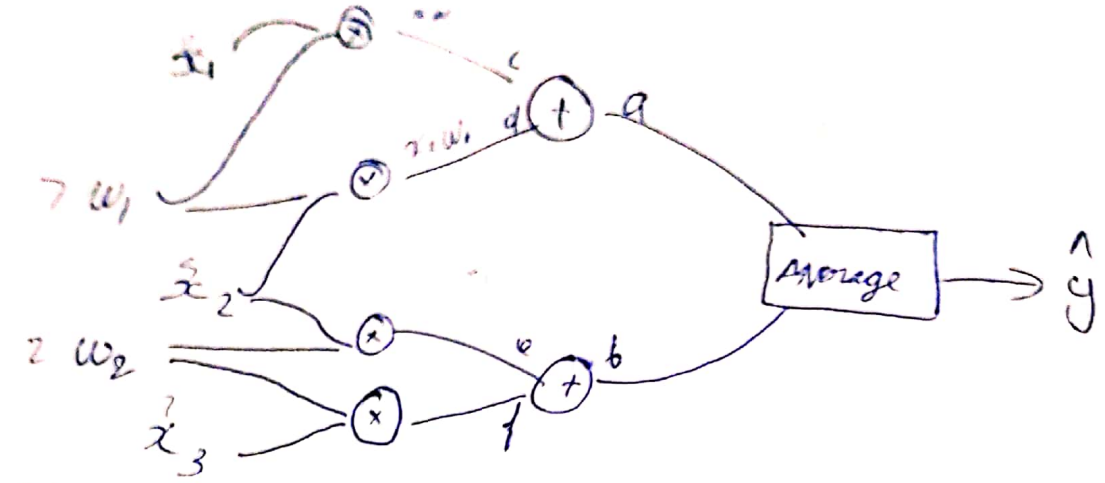
Majid Ahmad 191-1796 AI-J



01

target = 27

conv layer = [7, 2]

[25,7]

[5,6,1]

$[7 \times 2 + 2 \times 5 \;,\; 7 \times 5 + 2 \times 7]$

$[24, 49]$ $\quad \frac{24 + 49}{2} = 36.5$

$7 \times 3 + 2 \times 6, \; 7 \times 6 + 2 \times 1]$

33 44 $\quad \frac{33 + 44}{2} = 38.5$

Cost $= \frac{1}{n} \sum_i^n (y_i - t_i)^2$

$= \frac{1}{2} [(36.5 - 27)^2 + (38.5 - 27)^2]$

111.25

Derivative $\frac{1}{2} \sum_i^n (2 (y_i - t_i)$

$\frac{1}{2} ((2(36.5 - 27)^2 + 2(38.5 - 27))]$

$= 21$

Scanned with CamScanner

Avg derivative $= \frac{1}{2}(10)$

$\frac{1}{2}(21) = 10.5$ for a and b

$$\frac{\partial c}{\partial w_1} = \frac{2+3}{2} \Rightarrow 2.5 \times 10.5 \Rightarrow 26.25$$

$$\frac{\partial d}{\partial w_1} = \frac{5+6}{2} \Rightarrow 5.5 \times 10.5 \Rightarrow 57.75$$

$$\frac{\partial e}{\partial w_2} = \frac{5+6}{2} \Rightarrow 5.5 \times 10.5 \Rightarrow 57.75$$

$$\frac{\partial f}{\partial w_2} = \frac{7+1}{2} \Rightarrow 4 \times 10.5 \Rightarrow 42$$

$$\frac{\partial L}{\partial w_1} = 26.25 + 57.75 = 84$$

$$\frac{\partial L}{\partial w_2} = 57.75 + 42 = 99.75$$

Updating

$b_1 = 0.8$ $\quad b_1 = 0 + 84(1 - 0.8) = 16.8$

$b_2 = 0.999$ $\quad b_2 = 0 + 84^2(1 - 0.999) = 7.056$

based on online

$$w_1' = w_1 + \alpha \frac{S_1}{\sqrt{S_2}} \Rightarrow 7 - 0.01 \times \frac{16.8}{\sqrt{7.056}} = 6.9$$

$0 + 99.75(1 - 0.8) = 19.95$

$0 + 99.75^2(1 - 0.999) = 9.95$

$$w_2' = 2 - 0.01 \times \frac{19.95}{\sqrt{9.9}} = 1.9$$

Scanned with CamScanner

Iter 2 (3)

Conv be [6.937, 1.94] y = 63

[2 4 6]
[3 4 6]

[6.937 × 2 + 1.94 × 4, 6.937 × 4 + 1.94 × 6]

[6.937 × 3 + 1.94 × 2, 6937 × 2 + 1.94 × 6]

[21.634, 39.38] = 59.12 → 30.5

[24.69, 25.574] = $\frac{46.68}{2}$ ⇒ 25.10

Cost = $\frac{1}{2}\left[(25.10 - 63)^2 + (30.5 - 63)^2\right] = 979.6$

$= \frac{1}{2}\left(2(25.1 - 63) + 2(30.5 - 63)\right) = -70.4$

$\frac{\partial l}{\partial a} = \frac{\partial l}{\partial b} = -\frac{70}{2} = 35.22$

$\frac{\partial c}{\partial w_1} = \frac{2+3}{2} \Rightarrow 2.5 \times 35.22 = 88$

$\frac{\partial d}{\partial w_1} = \frac{4+2}{2} \Rightarrow 3 \times 35.22 = -105.6$

$\frac{\partial e}{\partial w_2} = \frac{4+2}{2} \Rightarrow 3 \times 35.22 = -105.6$

$\frac{\partial f}{\partial w_2} = \frac{6+6}{2} \Rightarrow 6 \times 35.22 = -211.2$

$\frac{\partial l}{\partial w_1} = 88 + 105.6 = -193.6$

$\frac{\partial l}{\partial w_2} = -211.6 + 105.6 = -317.2$

$S_1 = 16.8 \times 0.8 + (-193.6) \times (1 - 0.8) = -25.3$

$S_2 = 7.056 \times 0.99 + (+193.2) \times (1 - 0.999) = 37.5$

$w_1 = 6.937 - 0.01\left(\frac{-25.3}{\sqrt{37.5}}\right) = 6.98$

Scanned with CamScanner

w_2

$S_1 = 19.95 \times 0.8 + (1-0.8)(-317.2) = -47.5$

$S_2 = 9.95 \times 0.999 + (1-0.999)(-317.2)^2 = 110.3$

$w_2' = 1.94 - 0.01\left(\frac{-47.5}{\sqrt{110.3}}\right)$

$= 1.98$

Momentum ~~B~~ Optimizer

Same as before just here decay rate

$= 0.9$

$\partial L / \partial w = 84$

$\partial L / \partial w_2 = 99.75$

$(0.01 \times 84) + (0 \times 0.9)$

$= 0.84$

$w_1' \Rightarrow 7 - 0.84 = 6.16$

$-0.01 \times 99.75 + 0$

0.9975

$w_2' = 2 - 0.9975 = 1.0025$

$\begin{bmatrix} 2 & 4 & 6 \\ 3 & 2 & 6 \end{bmatrix}$ conv $[6.16 \quad 1.0025]$ $63 = y$

$[6.16 \times 2 + 1.0025 \times 4 \;,\; 6.16 \times 4 + 1.0025 \times 6]$

$\frac{[16.63 + 30.655]}{2} = 23.5$

$3 \times 6.16 + 2 \times 1.0025 \quad 2 \times 6.16 + 6 \times 1.0025$

$\frac{[20.5 + 18.34]}{2} = 19.4$

Scanned with CamScanner

Last

$$\frac{1}{2}\left(2(23.5-63)+2(19.4-63)\right)$$

$$= -83.1$$

$$\frac{\partial E}{\partial b} = \frac{\partial E}{\partial a} = -41.5$$

$$\frac{\partial E}{\partial w_1} = \frac{2+3}{2} = 2.5 \qquad \frac{\partial E}{\partial w_1 \partial w_2} = \frac{4+2}{2} = 3 \qquad \frac{\partial E}{\partial w_2} = \frac{6+6}{8} = 3$$

$w_1 \Rightarrow -41.5 \times 2.5 \Rightarrow -103.75 \rightarrow + = -228.25$

$w_1 \Rightarrow -41.5 \times 3 \Rightarrow -124.5$

$w_2 = -41.5 \times 8 \Rightarrow -124.5 \rightarrow + -373.5$

$w_2 = -41.5 \times 6 \Rightarrow -249$

$$-0.01(-228.25) + (0.84)(0.9)$$

$$= 1.53$$

$$w_1' = 6.16 + 1.53$$

$$w_1' = 7.69$$

$$\text{delta} = -0.01(\quad -373.5) + (0.9475)(\quad)^{0.9}$$

$$=$$

$$w_2' = 1.0025 + 2.84$$

$$w_2' = 3.84$$

Scanned with CamScanner