

Adesso :

In matrice ha pw : $w_{kj} =$

MATRICE DI

$\delta_k \cdot o_j$
PER GESTIONE DEL
BATCH

$$\begin{bmatrix} b_{row} \\ [L] \\ [..] \\ [..] \end{bmatrix}$$

x inizializzata con

, bins_row = np. ones (

(L, x.shape[L]))

aggiungiamo matrice di $\Delta_p w_{kj} = \delta_k \cdot o_j$:

$$\text{done } w_{kj}^{\text{new}} = w_{kj}^{\text{old}} + \gamma \cdot \Delta_p w_{kj}$$

$$\text{e batch } w_{kj}^{\text{new}} = w_{kj}^{\text{old}} + \gamma \cdot \sum_p \Delta_p w_{kj}$$

Questa matrice:

- viene reinizializzata ad ogni epoch
- deve avere la stessa shape di w_{kj}

per poi gestire nella classe trainer:

online

$$\left[\begin{array}{c} w_{kj} \end{array} \right] += \gamma \cdot \left[\begin{array}{c} \Delta_p w_{kj} \end{array} \right]$$

batch

$$\left[\begin{array}{c} w_{kj} \end{array} \right] += \gamma \cdot \left[\begin{array}{c} \sum_p \Delta_p w_{kj} \end{array} \right]$$

mini b?

IL BIAS IN MATRICE Δ_{Wkj}

$$\Delta_{Wkj} = \begin{bmatrix} [] & j_1 \\ [] & j_2 \\ \vdots & \vdots \\ [] & j_n \end{bmatrix} \rightarrow \text{bias} = \delta_k$$

non bias = $\delta_k \cdot o_j$

nel codice la matrice di Δ_{Wkj} non bias:

$$\Delta_{Wkj_non_b} = \Delta_{Wkj_w} =$$

$$\Delta_{Wkj} = \begin{bmatrix} \delta_{k_1} \cdot o_{j_1}, \dots \\ \vdots \\ \delta_{k_n} \cdot o_{j_n}, \dots \end{bmatrix} \quad (\pm)$$

quindi ogni elemento di
x-pur moltiplica ogni elem.
di δ_k , uso np. outer

$$x \quad (\times)$$

quindi esempio

$$\text{np. outer } (x_{-j} = [a, b], \delta_k = [z, z]) =$$

$$= \underbrace{x_{-j}[0] \cdot \delta_k[0]}_{\begin{bmatrix} a^2, ab \end{bmatrix}} \quad \underbrace{x_{-j}[1] \cdot \delta_k[1]}_{\begin{bmatrix} b, bb \end{bmatrix}}$$

Per aggiungere il bias posso aggiungere in testa un
1 a x_{prev} :

np. avrò ($x_{-j} = [1, a, b]$, $\delta_k = [l, r]$) =

$$x_{-j}[\phi] \cdot \delta_k[\phi] = \delta_k[\phi]$$

$$\left[\begin{array}{c} 1, \dots, \dots \\ \text{---} \\ \text{---} \end{array} \right] \cdot \left[\begin{array}{c} l, r \\ \text{---} \\ \text{---} \end{array} \right]$$

Una volta che Δw_{kj} posso aggiornare w_{kj}