CW (Autoenkoder)

Wejścia (input): $x^{(\alpha)} = (x_i^{(\alpha)})_{i=1}^2 = (x_1^{(\alpha)}, x_2^{(\alpha)}) \in \mathbb{R}^2 \quad (1 \le \alpha \le 2)$
$\square = 0.0$, $\blacksquare = 1.0$

$$x^{(1)} = (x_i^{(1)})_{i=1}^2 = (x_1^{(1)}, x_2^{(1)}) = \left|\begin{array}{|c|c|}\hline \blacksquare & \square \\ \hline\end{array}\right| \in \mathbb{R}^2$$
$$x^{(2)} = (x_i^{(2)})_{i=1}^2 = (x_1^{(2)}, x_2^{(2)}) = \left|\begin{array}{|c|c|}\hline \square & \blacksquare \\ \hline\end{array}\right| \in \mathbb{R}^2$$

(Obrazy z tylko 2 pikselami!)

<u>Enkoder</u> $x^{(\alpha)} \mapsto y^{(\alpha)} \quad (1 \le \alpha \le 2)$
Warstwa środkowa: $y^{(\alpha)} = (y_i^{(\alpha)})_{i=1}^1 = (y_1^{(\alpha)}) \in \mathbb{R}^1 \quad (1 \le \alpha \le 2)$

$$y_i^{(\alpha)} = f(\sum_{j=1}^{2} w_{ij} x_j^{(\alpha)} + b_i) \quad (1 \le i \le 1) \quad (1 \le \alpha \le 2)$$

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

<u>Dekoder</u> $y^{(\alpha)} \mapsto x'^{(\alpha)} \quad (1 \le \alpha \le 2)$
Wyjścia (output): $x'^{(\alpha)} = (x_i'^{(\alpha)})_{i=1}^2 = (x_1'^{(\alpha)}, x_2'^{(\alpha)}) \in \mathbb{R}^2 \quad (1 \le \alpha \le 2)$

$$x_k'^{(\alpha)} = f(\sum_{i=1}^{1} w'_{ki} y_i^{(\alpha)} + b'_k) \quad (1 \le k \le 2) \quad (1 \le \alpha \le 2)$$

Wyjścia (output) <u>dla wyświetlenia</u>: $x''^{(\alpha)} = (x_i''^{(\alpha)})_{i=1}^2 = (x_1''^{(\alpha)}, x_2''^{(\alpha)}) \in \{0,1\}^2 \quad (1 \le \alpha \le 2)$

$$x_k''^{(\alpha)} = f_1(\sum_{i=1}^{1} w'_{ki} y_i^{(\alpha)} + b'_k) \quad (1 \le k \le 2) \quad (1 \le \alpha \le 2)$$

$$f_1(x) = \begin{cases} 0 & \text{gdy } x < 0 \\ 1 & \text{gdy } x \ge 0 \end{cases}$$

Cel. Za pomocą metody gradientu

$$E = \frac{1}{2} \sum_{\alpha=1}^{2} \sum_{k=1}^{2} (x_k'^{(\alpha)} - x_k^{(\alpha)})^2 \to \text{minimum (lokalne).}$$

<u>Gradienty</u>

$$\frac{\partial E}{\partial w'_{pq}} = \sum_{\alpha=1}^{2} (x_p'^{(\alpha)} - x_p^{(\alpha)}) f'(\sum_{i=1}^{1} w'_{pi} y_i^{(\alpha)} + b'_p) y_q^{(\alpha)} \quad (1 \le p \le 2, 1 \le q \le 1)$$

$$\frac{\partial E}{\partial b'_p} = \sum_{\alpha=1}^{2} (x_p'^{(\alpha)} - x_p^{(\alpha)}) f'(\sum_{i=1}^{1} w'_{pi} y_i^{(\alpha)} + b'_p) \quad (1 \le p \le 2)$$

$$\frac{\partial E}{\partial w_{pq}} = \sum_{\alpha=1}^{2} \sum_{k=1}^{2} (x_k'^{(\alpha)} - x_k^{(\alpha)}) f'(\sum_{i=1}^{1} w'_{ki} y_i^{(\alpha)} + b'_k) w'_{kp} f'(\sum_{j=1}^{2} w_{pj} x_j^{(\alpha)} + b_p) x_q^{(\alpha)} \quad (1 \le p \le 1, 1 \le q \le 2)$$

$$\frac{\partial E}{\partial b_p}=\sum_{\alpha=1}^{2}\sum_{k=1}^{2}(x_k'^{(\alpha)}-x_k^{(\alpha)})f'(\sum_{i=1}^{1}w'_{ki}y_i^{(\alpha)}+b'_k)w'_{kp}f'(\sum_{j=1}^{2}w_{pj}x_j^{(\alpha)}+b_p)\quad (1\le p\le 1)$$

Zadanie. Implementować algorytm metody gradientu dla autoenkodera.

Zad. (1) Wybrać różne c, ε i β. Dobrać warunek początkowy $w_{ij}^{\mathrm{old}}, b_i^{\mathrm{old}}, w_{ij}'^{\mathrm{old}}, b_i'^{\mathrm{old}}$.

Wskazówki do dobrania warunku początkowego.
Enkoder
$x^{(1)}=(1,0)\mapsto y_1^{(1)}=f(w_{11}+b_1)\approx 0$ $(\alpha=1)$
$x^{(2)}=(0,1)\mapsto y_1^{(2)}=f(w_{12}+b_1)\approx 1$ $(\alpha=2)$
Dekoder
$y^{(1)}\mapsto x'^{(1)}=(f(w'_{11}y_1^{(1)}+b'_1), f(w'_{21}y_1^{(1)}+b'_2))\approx(f(b'_1),f(b'_2))\approx(1,0)$ $(\alpha=1)$
$y^{(2)}\mapsto x'^{(2)}=(f(w'_{11}y_1^{(2)}+b'_1), f(w'_{21}y_1^{(2)}+b'_2))\approx(f(w'_{11}+b'_1),f(w'_{21}+b'_2))\approx(0,1)$ $(\alpha=2)$
Wnioski do dobrania warunków początkowych:
Dowrać warunki początkowe aby spełnić następujące warunki.
$f(w_{11}+b_1)\approx 0,\ f(w_{12}+b_1)\approx 1;\ f(b'_1)\approx 1,\ f(b'_2)\approx 0;\ f(w'_{11}+b'_1)\approx 0,\ f(w'_{21}+b'_2))\approx 1$

Po zakończeniu iteracji metody gradientu wyświetlić **obrazy** $x''^{(\alpha)}$ $(1\le\alpha\le 2)$. (Uwaga: $x''^{(\alpha)}$ powinien być $x^{(\alpha)}$ $(1\le\alpha\le 2)$).
Zaznaczyć jakie c, ε, β i warunek początkowy $w_{ij}^{\mathrm{old}}, b_i^{\mathrm{old}}, w_{ij}'^{\mathrm{old}}, b_i'^{\mathrm{old}}$ były wybrane.

Zad. (2) Wybrać różne c, ε i β. Ustalić N i losować warunek początkowy $w_{ij}^{\mathrm{old}}, b_i^{\mathrm{old}}, w_{ij}'^{\mathrm{old}}, b_i'^{\mathrm{old}}\in[-N,N]$. Po zakończeniu iteracji metody gradientu wyświetlić **obrazy** $x''^{(\alpha)}$ $(1\le\alpha\le 2)$. (Uwaga: $x''^{(\alpha)}$ powinien być $x^{(\alpha)}$ $(1\le\alpha\le 2)$).
Próbować różne parametry c, ε, β oraz N i zbadać zależność od tych parametrów.

Notacja. (Propozycja)
$x_j^{(\alpha)}\rightsquigarrow x[a][j],\ y_i^{(\alpha)}\rightsquigarrow y[a][i]\ (1\le i\le 1!),\ x_k'^{(\alpha)}\rightsquigarrow xp[a][k],\quad x_k''^{(\alpha)}\rightsquigarrow xpp[a][k]$
$w_{ij}^{\mathrm{old}}\rightsquigarrow w\_old[i][j],\ w_{ij}^{\mathrm{new}}\rightsquigarrow w\_new[i][j],\ b_i^{\mathrm{old}}\rightsquigarrow b\_old[i],\ b_i^{\mathrm{new}}\rightsquigarrow b\_new[i]$
$w_{ki}'^{\mathrm{old}}\rightsquigarrow w\_pold[k][i],\ w_{ki}'^{\mathrm{new}}\rightsquigarrow w\_pnew[k][i],\ b_k'^{\mathrm{old}}\rightsquigarrow b\_pold[k],\ b_k'^{\mathrm{new}}\rightsquigarrow b\_pnew[k]$
$\frac{\partial E}{\partial w'_{pq}}\rightsquigarrow DE\_wp[p][q], \frac{\partial E}{\partial b'_p}\rightsquigarrow DE\_bp[p], \frac{\partial E}{w_{pq}}\rightsquigarrow DE\_w[p][q], \frac{\partial E}{\partial b_p}\rightsquigarrow DE\_b[p]$
$f\rightsquigarrow f,\ f_1\rightsquigarrow f_1,\ f'\rightsquigarrow Df,\ e^{-\beta x}\rightsquigarrow math.exp((-1)*beta*x)$ (Python?)
$c\rightsquigarrow c,\ \varepsilon\rightsquigarrow epsilon$ (lub ε w Python?), $\beta\rightsquigarrow beta$ (lub β w Python?), $N\rightsquigarrow N$