Homework 1

Arsene Andreea

$x = [1, 3, 0]$

$$w = \begin{bmatrix} 0.3 & 0.1 & -2 \\ -0.6 & -0.5 & 2 \\ -1 & -0.5 & 0.1 \end{bmatrix}$$

$b = [0.1 \ 0.1 \ 0.1]$

$y = [0, 1, 0]$

$$z = W^T x + b = \begin{bmatrix} 0.3 & 0.1 & -2 \\ -0.6 & -0.5 & 2 \\ -1 & -0.5 & 0.1 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix} + \begin{bmatrix} 0.1 \\ 0.1 \\ 0.1 \end{bmatrix}$$

$z_0 = [0.3 \ 0.1 \ -2][1 \ 3 \ 0] + 0.1 = 0.3 + 0.3 + 0.1 = 0.7$

$z_1 = [-0.6 \ -0.5 \ 2][1 \ 3 \ 0] + 0.1 = -0.6 - 1,5 + 0.1 = -2$

$z_2 = [-1 \ -0.5 \ 0.1][1 \ 3 \ 0] + 0.1 = -1 - 1,5 + 0.1 = -2,4$

$$\hat{y}_i = \frac{e^{z_i}}{\sum_j e^{z_j}} \quad i = 0 \dots n \qquad \sum_j e^{z_j} = e^{0.7} + e^{-2} + e^{-2,4} = 2.01 + 0.13 + 0.09 = 2.23$$

$$\hat{y}_0 = \frac{e^{0.7}}{2.23} = \frac{2.01}{2.23} = 0.90 \qquad \hat{y}_2 = \frac{0.09}{2.23} = 0.04$$

$$\hat{y}_1 = \frac{0.13}{2.23} = 0.05$$

$\nabla_z L = \hat{y} - y = [0.90 - 0 \quad 0.05 - 1 \quad 0.04 - 0] = [0,90 \ -0,95 \ 0.04]$

$$\nabla_W L = \nabla_z L^T x = \begin{bmatrix} 0,90 \\ -0,95 \\ 0.04 \end{bmatrix} [1 \ 3 \ 0] = \begin{bmatrix} 0.90 & 2,7 & 0 \\ -0,95 & -2,85 & 0 \\ 0.04 & 0,12 & 0 \end{bmatrix}$$

$\nabla_b L = \nabla_z L = [0,90 \ -0,95 \ 0.04]$

$\eta = 0.1$

$$w \leftarrow w - \eta \nabla_W L = W - 0.1 \begin{bmatrix} 0.90 & 2.7 & 0 \\ -0.95 & -2.85 & 0 \\ 0.04 & 0.12 & 0 \end{bmatrix} = W - \begin{bmatrix} 0.09 & 0.27 & 0 \\ -0.095 & -0.285 & 0 \\ 0.004 & 0.012 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 0.21 & -0,17 & -2 \\ -0.50 & -0,215 & 2 \\ -1.004 & -0,51 & 0.1 \end{bmatrix}$$

$b \leftarrow b - \eta \nabla_b L = [0.1, 0.1, 0.1] - 0.1 [0.90 \ -0.95 \ 0.04]$

$= b - [0.09 \ -0.095 \ 0.004] = [0.01, 0.195, 0.096]$