$X = (1, 3, 0)$

$Y = (0, 1, 0)$; "$Y = 1$"

$$W = \begin{pmatrix} 0,1 & 0,2 & -0,3 \\ -0,6 & -0,5 & 2 \\ -0,2 & 0,5 & -3 \end{pmatrix}$$

$b = (0,1 \quad 0,1 \quad 0,1)$

$$\mathrm{softmax}(z_i) = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}}$$

$$\hat{Y} = \mathrm{softmax}(W \cdot X + b)$$

$$C = -\frac{1}{n} \cdot \sum_{X} [Y \ln \hat{Y} + (1 - Y) \cdot \ln(1 - \hat{Y})]$$

Cross entropie pentru cazul de output binar

$\Downarrow$ mai generalizat

$$C = -\frac{1}{n} \sum_{X} \sum_{j} [Y_j \cdot \ln \hat{Y}_j + (1 - Y_j) \cdot \ln(1 - \hat{Y}_j)]$$

Scanat cu CamScanner

Vrem să calculăm $\frac{\partial C}{\partial w}$ și $\frac{\partial C}{\partial b}$ pentru a minimiza C

Vom considera activarea $\sigma(z_i) = \frac{1}{1+e^{-z_i}}$ pentru a simplifica calculele

Aflăm prima dată $\frac{\partial C}{\partial z_i}$;

$$\frac{\partial C}{\partial z_i} = \frac{\partial C}{\partial \hat{Y}_i} \cdot \frac{\partial \hat{Y}_i}{\partial z_i}$$

$$\frac{\partial C}{\partial \hat{Y}_i} = -\frac{1}{m} \cdot \sum_x Y_i \cdot \frac{1}{\hat{Y}_i} + (1-Y_i) \cdot \frac{1}{(1-\hat{Y}_i)} \cdot (1-\hat{Y}_i)'$$

$$\Rightarrow \boxed{\frac{\partial C}{\partial \hat{Y}_i} = -\frac{1}{m} \cdot \sum_x \frac{Y_i}{\hat{Y}_i} - \frac{(1-Y_i)}{(1-\hat{Y}_i)}}$$

$$\frac{\partial \hat{Y}_i}{\partial z_i} = \left(\frac{1}{1+e^{-z_i}}\right)' = \left(\left(1+e^{-z_i}\right)^{-1}\right)' = -(1+e^{-z_i})^{-2} \cdot (1+e^{-z_i})'$$

$$= -\frac{1}{(1+e^{-z_i})^2} \cdot e^{-z_i} \cdot (-z_i)' = \frac{e^{-z_i}}{(1+e^{-z_i})^2}$$

$$\Leftrightarrow \frac{\partial \hat{Y}_i}{\partial z_i} = \frac{1}{1+e^{-z_i}} \cdot \frac{e^{-z_i}}{1+e^{-z_i}} \Leftrightarrow \frac{\partial \hat{Y}_i}{\partial z_i} = \hat{Y}_i \cdot \left(\frac{1+e^{-z_i}}{1+e^{-z_i}} - \frac{1}{1+e^{-z_i}}\right)$$

$$= \hat{Y}_i \cdot (1-\hat{Y}_i) \Leftrightarrow \boxed{\frac{\partial \hat{Y}_i}{\partial z_i} = \hat{Y}_i \cdot (1-\hat{Y}_i)}$$

$(1-\hat{Y}_i), \quad \hat{Y}_i$

$$\frac{\partial C}{\partial z_i} = \left[-\frac{1}{m} \cdot \sum_x \frac{Y_i}{\hat{Y}_i} - \frac{(1-Y_i)}{(1-\hat{Y}_i)} \right] \cdot \hat{Y}_i \cdot (1-\hat{Y}_i)$$

$$\Rightarrow \frac{\partial C}{\partial z_i} = \left[-\frac{1}{m} \cdot \sum_x \frac{Y_i \cdot (1-\hat{Y}_i) - \hat{Y}_i \cdot (1-Y_i)}{\hat{Y}_i \cdot (1-\hat{Y}_i)} \right] \cdot \hat{Y}_i \cdot (1-\hat{Y}_i)$$

$$\Rightarrow \frac{\partial C}{\partial z_i} = -\frac{1}{m} \cdot \sum_x Y_i - \cancel{Y_i \cdot \hat{Y}_i} - \hat{Y}_i + \cancel{\hat{Y}_i \cdot Y_i}$$

$$\Rightarrow \frac{\partial C}{\partial z_i} = -\frac{1}{m} \cdot \sum_x Y_i - \hat{Y}_i$$

$$\Rightarrow \boxed{\frac{\partial C}{\partial z_i} = \frac{1}{m} \cdot \sum_x \hat{Y}_i - Y_i}$$

$$\frac{\partial C}{\partial w} = \frac{\partial C}{\partial z_i} \cdot \frac{\partial C}{\partial w} \; ; \quad \frac{\partial C}{\partial w} = (w \cdot x + b)' = x \Rightarrow \frac{\partial C}{\partial w} = x$$

$$\Rightarrow \boxed{\frac{\partial C}{\partial w} = \frac{1}{m} \cdot \sum_x (\hat{Y}_i - Y_i) \cdot x}$$

$$\frac{\partial C}{\partial b} = \frac{\partial C}{\partial z_i} \cdot \frac{\partial C}{\partial b} \; ; \quad \frac{\partial C}{\partial b} = (w \cdot x + b)' = 1 \Rightarrow \frac{\partial C}{\partial b} = 1$$

$$\Rightarrow \boxed{\frac{\partial C}{\partial b} = \frac{1}{m} \cdot \sum_x (\hat{Y}_i - Y_i)}$$

Pentru activarea softmax și funcția de cost Cross entropie:

$$\boxed{\frac{\partial C}{\partial z_i} = \frac{1}{n} \cdot \sum_x \hat{Y}_i - Y_i}$$

=> putem folosi rezultatele obținute anterior

Beed Forward

$$z = w \cdot x + b$$

$$Z = \begin{pmatrix} 0,1 & 0,2 & -0,3 \\ -0,6 & -0,5 & 2 \\ -0,2 & 0,5 & -3 \end{pmatrix} \cdot (1 \quad 3 \quad 0) +$$

$$(0,1 \quad 0,1 \quad 0,1)$$

$$\Leftrightarrow z = (0,7 \quad -2,1 \quad 1,3) + (0,1 \quad 0,1 \quad 0,1)$$

$$\Leftrightarrow z = (0,8 \quad -2 \quad 1,4)$$

$$\hat{Y} = softmax(z) \Leftrightarrow \hat{Y} = (0,34 \quad 0,21 \quad 0,63)$$

Calcularea gradienților

$$\frac{\partial C}{\partial w} = \frac{1}{m} \cdot \sum_{x} (\hat{Y}_i - Y_i) \cdot x \Rightarrow \frac{\partial C}{\partial w} = \{[(0,34 \quad 0,21 \quad 0,63) - (0 \quad 1 \quad 0)] \cdot (1 \quad 3 \quad 0)\}^t$$

$$\Rightarrow \frac{\partial C}{\partial w} = [(0,34 \quad -0,97 \quad 0,63) \cdot (1 \quad 3 \quad 0)]^t$$

$$\Rightarrow \boxed{\frac{\partial C}{\partial w} = \begin{pmatrix} 0,34 & -0,97 & 0,63 \\ 1,04 & -2,93 & 1,89 \\ 0 & 0 & 0 \end{pmatrix}}$$

$$\frac{\partial C}{\partial b} = \frac{1}{m} \cdot \sum_{x} (\hat{Y}_i - Y_i) = (0,34 \quad 0,21 \quad 0,63) - (0 \quad 1 \quad 0)$$

$$\Rightarrow \boxed{\frac{\partial C}{\partial b} = (0,34 \quad -0,97 \quad 0,63)}$$

Actualizarea Weights și Biases

$$w \leftarrow w - \frac{\partial C}{\partial w}$$

$$w \leftarrow \begin{pmatrix} 0,1 & 0,2 & -0,3 \\ -0,6 & -0,5 & 2 \\ -0,2 & 0,5 & -3 \end{pmatrix} - \begin{pmatrix} 0,34 & -0,97 & 0,63 \\ 1,04 & -2,93 & 1,89 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\boxed{w \leftarrow \begin{pmatrix} -0,24 & 1,17 & -0,93 \\ -1,64 & 2,43 & 0,1 \\ -0,2 & 0,5 & -3 \end{pmatrix}}$$

$$b \leftarrow b - \frac{\partial C}{\partial b}$$

$$b \leftarrow (0,1 \quad 0,1 \quad 0,1) - (0,34 \quad -0,97 \quad 0,63)$$

$$\boxed{b \leftarrow (-0,24 \quad 1,07 \quad -0,53)}$$