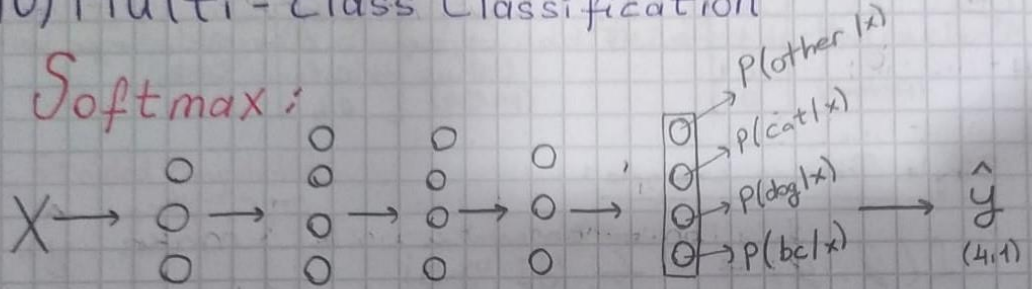
10) Multi-class Classification

Softmax:



Output layer'da multi sonuç üretmek için kullanılır.

Activation function:

Temporary variable $t = e^{(z^{[l]})}$

$$a^{[l]}_{(4,1)} = \frac{e^{(z^{[l]})}}{\sum_{i=1}^{4} t_i} \quad ; \quad a^{[l]}_i = \frac{t_i}{\sum_{i=1}^{4} t_i}$$

4 class olduğunu farz edelim.

ÖR

$$Z^{[l]} = \begin{bmatrix} 5 \\ 2 \\ -1 \\ 3 \end{bmatrix} \rightarrow t = \begin{bmatrix} e^5 \\ e^2 \\ e^{-1} \\ e^3 \end{bmatrix} = \begin{bmatrix} 148.4 \\ 7.4 \\ 0.4 \\ 20.1 \end{bmatrix} \quad ; \sum_{i=1}^{4} t_i = 176.3$$

$$a^{[l]} = \frac{t}{176.3} \rightarrow \begin{array}{l} \rightarrow 0.842 \text{ (other)} \\ \rightarrow 0.042 \text{ (cat)} \\ \rightarrow 0.002 \text{ (dog)} \\ \rightarrow 0.114 \text{ (bc)} \end{array}$$

* Yani; $a^{[l]}_{(4,1)} = g^{[l]}\left(z^{[l]}_{(4,1)}\right)$

Understanding Softmax: (Output layerda kullanilir)

$$z^{[l]} = \begin{bmatrix} 5 \\ 2 \\ -1 \\ 3 \end{bmatrix} \qquad t = \begin{bmatrix} e^5 \\ e^2 \\ e^{-1} \\ e^3 \end{bmatrix} \qquad C = 4 \text{ (number of classes)}$$

$$a^{[l]} = g^{[l]}\left(z^{[l]}\right) = \begin{bmatrix} e^5/(e^5+e^2+e^{-1}+e^3) \\ e^2/(e^5+e^2+e^{-1}+e^3) \\ e^{-1}/(e^5+e^2+e^{-1}+e^3) \\ e^3/(e^5+e^2+e^{-1}+e^3) \end{bmatrix} = \begin{bmatrix} 0.842 \\ 0.042 \\ 0.002 \\ 0.114 \end{bmatrix}$$

$$\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$ ← Biggest prob.

Loss Func.:

ÖR

C=4

$$y_{(\text{Actual})} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \rightarrow \text{cat} \qquad \hat{y}_{(\text{predict})} = \begin{bmatrix} 0.3 \\ 0.2 \\ 0.1 \\ 0.4 \end{bmatrix} \rightarrow \text{0.2 chance being a cat}$$

Single Training Example:

$$\mathcal{L}(\hat{y}, y) = -\sum_{j=1}^{4} y_j \log \hat{y}_j$$

$y_1 = y_3 = y_4 = 0$

$y_2 = 1$

} make $\hat{y}_2$ big to decrease loss

Entire Training set:

$$J(w^{[l]}, b^{[l]}, \ldots) = \frac{1}{m} \sum_{i=1}^{m} \mathcal{L}\left(\hat{y}^{(i)}, y^{(i)}\right)$$

$$Y = \left[y^{(1)}, y^{(2)}, y^{(3)} \ldots y^{(m)}\right] \qquad \hat{Y} = \left[\hat{y}^{(1)}, \ldots, \hat{y}^{(m)}\right]$$

$$= \begin{bmatrix} 0 & 0 & 1 & \\ 1 & 0 & 0 & \cdots \\ 0 & 1 & 0 & \\ 0 & 0 & 0 & \end{bmatrix} \qquad = \begin{bmatrix} 0.3 & & \\ 0.2 & \cdots & \\ 0.1 & & \\ 0.4 & & \end{bmatrix}$$

#cr(4,m) (4,m)