Yusuf GÜR

2532810

BackPropogation

Sigmoid

$$X_{out} = \sigma(X_{in}) = \frac{1}{1+e^{-X_{in}}}$$

$$\frac{\partial X_{out}}{\partial X_{in}} = (-1)(1+e^{-X_{in}})^{-2}(e^{-X_{in}})(-1) = \frac{e^{-X_{in}}}{(1+e^{-X_{in}})^2}$$

$$= \frac{1}{(1+e^{-X_{in}})} \cdot \frac{e^{-X_{in}}}{1+e^{-X_{in}}}$$

$$= \sigma(X_{in}) \cdot \frac{1+e^{-X_{in}}-1}{(1+e^{-X_{in}})}$$

$$= \sigma(X_{in}) \left(\frac{1+e^{-X_{in}}}{1+e^{-X_{in}}} - \frac{1}{1+e^{-X_{in}}} \right)$$

$$= \sigma(X_{in}) \cdot (1-\sigma(X_{in}))$$

tanh

$$X_{out} = \tanh(X_{in}) = \frac{\exp(2X_{in})-1}{\exp(2X_{in})+1} \rightarrow \left(\frac{\sin(x)}{\cos(x)}\right)'$$

$$\frac{\partial X_{out}}{\partial X_{in}} = \frac{(\sin(x_{in}))' \cdot \cos(x_{in}) - \sin(x_{in}) \cdot (\cos(x_{in}))'}{(\cos(x_{in}))^2}$$

$$= \frac{(\cos(x_{in}))^2 - (\sin(x_{in}))^2}{(\cos(x_{in}))^2}$$

$$= 1 - \left(\frac{\sin(x_{in})}{\cos(x_{in})}\right)^2$$

$$= 1 - \tanh(X_{in})^2$$

ReLU

$$X_{out} = \max(0, X_{in})$$

$$X_{out} = \begin{cases} X_{in}, & \text{if } x > 0 \\ 0, & \text{if } x \leq 0 \end{cases}$$

$$\frac{\partial X_{out}}{\partial X_{in}} = \begin{cases} 1, & \text{if } x > 0 \\ 0, & \text{if } x \leq 0 \end{cases}$$

Softmax

$$(X_{out})_i = \frac{e^{-\beta(X_{in})_i}}{\sum_k e^{-\beta(X_{in})_k}} \quad \begin{matrix} \rightarrow g(x) \\ \rightarrow h(x) \end{matrix}$$

$$\frac{\partial X_{out}}{\partial X_{in}} = \frac{g'(x) \cdot h(x) - g(x) \cdot h'(x)}{h(x)^2}$$

$$\frac{\partial g_i}{\partial x_j} = \begin{cases} e^{x_j}, & i=j \\ 0, & \text{else} \end{cases}$$

$$\frac{\partial h_i}{\partial x_j} = \frac{\partial(e^{x_1}+e^{x_2}+\dots+e^{x_n})}{\partial x_j} = e^{x_j}$$

if $i=j$

$$= \frac{e^{x_j} \cdot \sum_{k=1}^{n} e^{x_k} - e^{x_j} \cdot e^{x_i}}{\left(\sum_{k=1}^{n} e^{x_k}\right)^2}$$

$$= \frac{e^{x_i} \cdot \sum_{k=1}^{n} e^{x_k} - e^{x_j}}{\sum_{k=1}^{n} e^{x_k} \cdot \sum_{k=1}^{n} e^{x_k}}$$

$$= \frac{e^{x_i}}{\sum_{k=1}^{n} e^{x_k}} \left(\frac{\sum_{k=1}^{n} e^{x_k}}{\sum_{k=1}^{n} e^{x_k}} - \frac{e^{x_j}}{\sum_{k=1}^{n} e^{x_k}} \right)$$

$$= \sigma(x_i)(1-\sigma(x_j))$$

if $i \neq j$

$$= \frac{0 - e^{x_j} \cdot e^{x_i}}{\left(\sum_{k=1}^{n} e^{x_k}\right)^2}$$

$$= -\frac{e^{x_j}}{\sum_{k=1}^{n} e^{x_k}} \cdot \frac{e^{x_i}}{\sum_{k=1}^{n} e^{x_k}}$$

$$= -\sigma(x_j) \cdot \sigma(x_i)$$

$$\begin{cases} \sigma(x_i)(1-\sigma(x_j)), & i=j \\ -\sigma(x_j) \cdot \sigma(x_i), & i \neq j \end{cases}$$