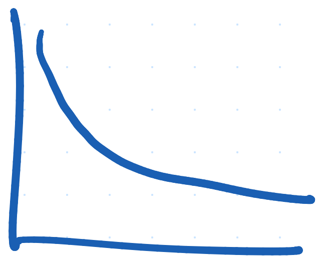
Lecture-2

25/01/2026

I/p → $\boxed{M}$ → o/p.



(Matrix)

Tensor.

$\boxed{C}$

$\boxed{C}$

Adam(w).

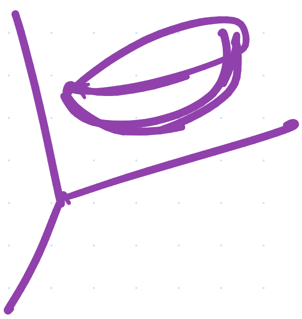
$\boxed{D}$

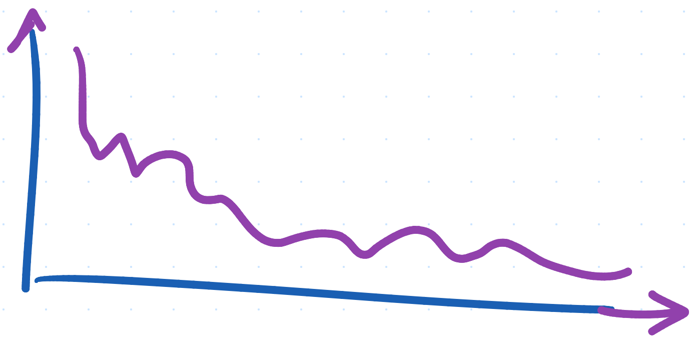
$\boxed{D-}$

O

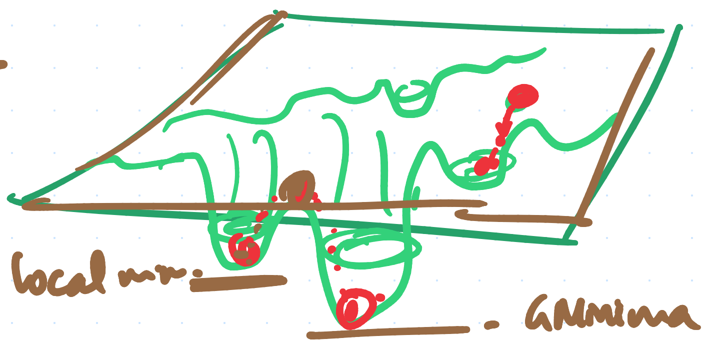
L

SGD.





SGD.



m. -

$\underline{v}^T \underline{v}$

$v = \begin{pmatrix} \cdot \\ \cdot \\ \cdot \end{pmatrix}$ $\quad v = \begin{pmatrix} \cdot & \cdot & \cdot \end{pmatrix}.$

$v = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$

$$\underbrace{\begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}}_{\text{inner.} \to \text{dot } \underline{\text{product}}.} = 1\cdot1 + 2\cdot2 + 3\cdot3 = 14$$

$$v . v^T = \underbrace{\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \end{pmatrix}}_{\text{outer} \to .} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{pmatrix}.$$

$$\begin{array}{ccc} H & W & C \\ 224 \times & 224 \times & 3 \\ \hline 0 & 1 & 2. \end{array}$$

permute $(2, 0, 1)$.

$3 \times 224 \times 224.$

$\boxed{0}\ \boxed{1}\ \boxed{2}$

$$\begin{array}{ccc} 2 & 0 & 1 \\ C & H & W. \end{array}$$

<u>Image.</u> $\leftarrow [B, C, H, W]$

$$y = 3 \cdot x^2 + 15$$

$$y\Big|_{x=2} = \frac{\partial y}{\partial x} = \underbrace{6x}_{\text{gradient.}} + 0.$$

$x = 3.$ $\qquad 6 \times 3 = \underline{\underline{18}}.$

$$\boxed{z = 5 \cdot x^3 + 7 \cdot y^2}$$

$$\frac{\partial z}{\partial x} \quad \& \quad \frac{\partial z}{\partial y}$$

$$\underline{\underline{\frac{\partial z}{\partial x}}} = 15 \cdot x^2 + 14 \cancel{\frac{\partial y}{\partial x}} \rightarrow 0. \qquad = 135.$$

$$\frac{\partial z}{\partial y} = \cancel{15x^2 \frac{\partial x}{\partial y}} + \underline{\underline{14y}}. \qquad = 28.$$

$x = 3.$

$$\begin{array}{r} 15 \\ \times 9 \\ \hline 135 \end{array}$$

$$\therefore \left.\frac{\partial z}{\partial x}\right|_{x=3} = 15(3)^2 = 15 \times 9 = \underline{135}.$$

$$y = \frac{1}{\ell(x)} \sum_i \left[(x_i + 2)^2 + 3 \right].$$

$$a_i = (x_i + 2) \Rightarrow a = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} \qquad x = \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} \begin{array}{l} \rightarrow x(1) \\ \rightarrow x(2) \\ \rightarrow x(3) \end{array}$$

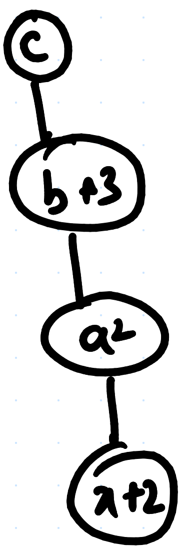
$$a = (x+2) \rightarrow \frac{\partial a}{\partial x} = 1.$$

$$b = a^2 \qquad \frac{\partial b}{\partial a} = 2a$$

$$c = b + 3. \longrightarrow \frac{\partial c}{\partial b} = 1$$

$$y = \underline{c}.\text{mean}() \rightarrow \frac{\partial y}{\partial c} = \frac{1}{3}.$$

$$\frac{\partial y}{\partial c_i} \cdot \frac{\partial c_i}{\partial b_i} \cdot \frac{\partial b_i}{\partial a_i} \cdot \frac{\partial a_i}{\partial x_i} \rightarrow \frac{\partial y}{\partial x_i}.$$



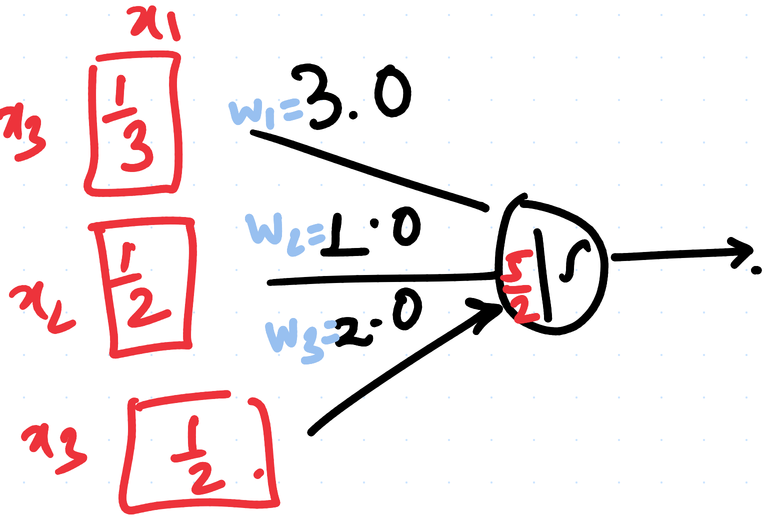
$$\Rightarrow \frac{1}{3} \times 1 \times 2.a_i \times 1.$$

$$\Rightarrow \frac{2}{3}[2, 3, 4] \longrightarrow \begin{bmatrix} \frac{4}{3}, 2, \frac{8}{3} \end{bmatrix}$$

gradient of y w.r.t. x

$$\sigma(x) = \frac{1}{1+e^{-x}}.$$

$w_1 . x_1 + w_2 . x_2 + w_3 . x_3$



$1 + \frac{1}{2} + 1$

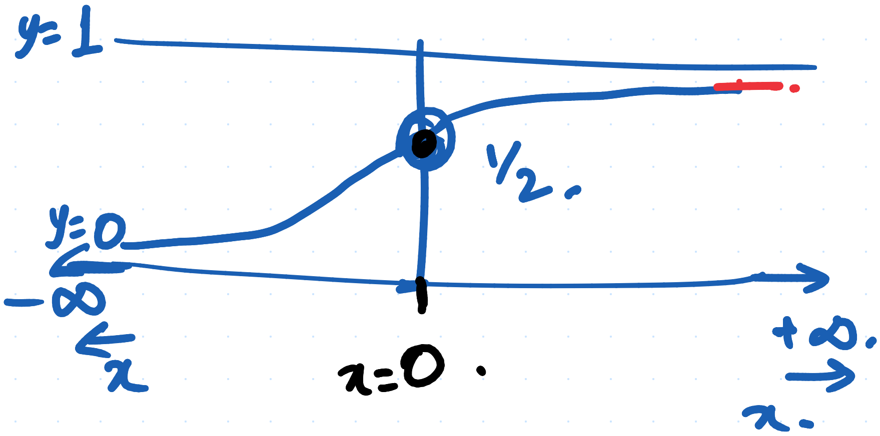
$\frac{5}{2}$.

$$O/p. = \frac{1}{1+e^{-5/2}}$$

$\frac{1}{1+e^{-x}}$.

when $x = 0$; $\sigma(x) = \frac{1}{2}$.

$\Rightarrow \frac{1}{1+e^{0}} = \frac{1}{2}$.



when $x = +\infty$.

$$\frac{1}{1+e^{-\infty}} = \frac{1}{1+\frac{1}{e^{\infty}}} = \frac{1}{1+\frac{1}{\infty}} = \frac{1}{1+0}$$

when $\boxed{x=\infty}$, $\boxed{\sigma(x)=1}$ $= 1.$

when $\boxed{x=-\infty}$. $\boxed{\sigma(x)=0}$.

$$\frac{1}{1+e^{-x}} = \frac{1}{1+e^{\infty}} = \frac{1}{1+\infty} = \frac{1}{\infty} = 0.$$

$$\boxed{\sigma(x) = \frac{1}{1+e^{-x}}}$$

$$\frac{\partial}{\partial x}\sigma(x) = ?$$

Derivative of Sigmoid

$$\frac{\partial}{\partial x}\left(\underline{\log_e(\sigma(x))}\right) = \frac{\partial}{\partial x}\left(\underline{\log_e(1+e^{-x})^{-1}}\right).$$

$$\Rightarrow \frac{\sigma'(x)}{\sigma(x)} = +\frac{1}{(1+e^{-x})}e^{-x}$$

$$\Rightarrow \sigma'(x). = \sigma(x). \frac{e^{-x}}{1+e^{-x}}$$

$$\boxed{\sigma'(x) = \sigma(x)(1-\sigma(x)).}$$

$$\frac{\cancel{1+}e^{-x}\cancel{+1}}{1+e^{-x}} = \left(1 - \frac{1}{1+e^{-x}}\right) = \frac{e^{-x}}{1+e^{-x}}$$

$$\sigma'(x) = \sigma(x)\left(1-\sigma(x)\right).$$

Maximum value of derivative, i.e, $\sigma'(x) = ?$.

$$z = x(1-x) = x - x^2$$

$$\frac{\partial z}{\partial x} = 1 - 2x$$

$$1 - 2x = 0.$$

$$\therefore \boxed{x = \frac{1}{2}}$$

$$\sigma'(x) = \sigma(x)\left(1-\sigma(x)\right) = \frac{1}{2}\left(1 - \frac{1}{2}\right) = \frac{1}{2} \times \frac{1}{2}$$

Max value.

$$= \frac{1}{4}.$$

$$\text{ReLU} = \max(0, x).$$

$$\boxed{\text{ReLU} = \begin{cases} 0 & \text{if } x < 0. \\ x & \text{if } x \geq 0. \end{cases}}$$

$$\boxed{\tanh = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}}}$$

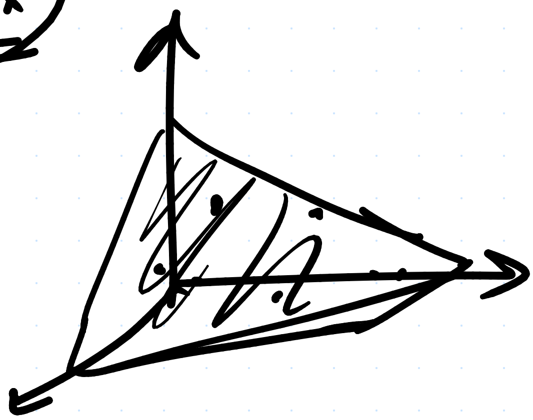
$$\text{Softmax.} = \frac{e^{x_i}}{\sum_{x} e^{x_i}}$$

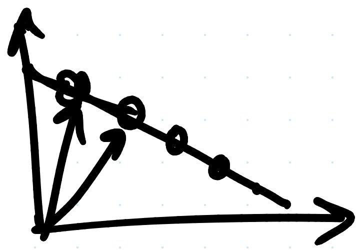
$$\begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} \xrightarrow{\text{Softmax.}} \left\{ \underbrace{\frac{e^0}{e^0+e^1+e^2}, \frac{e^1}{e^0+e^1+e^2}, \frac{e^2}{\ldots}}_{} \right\}$$

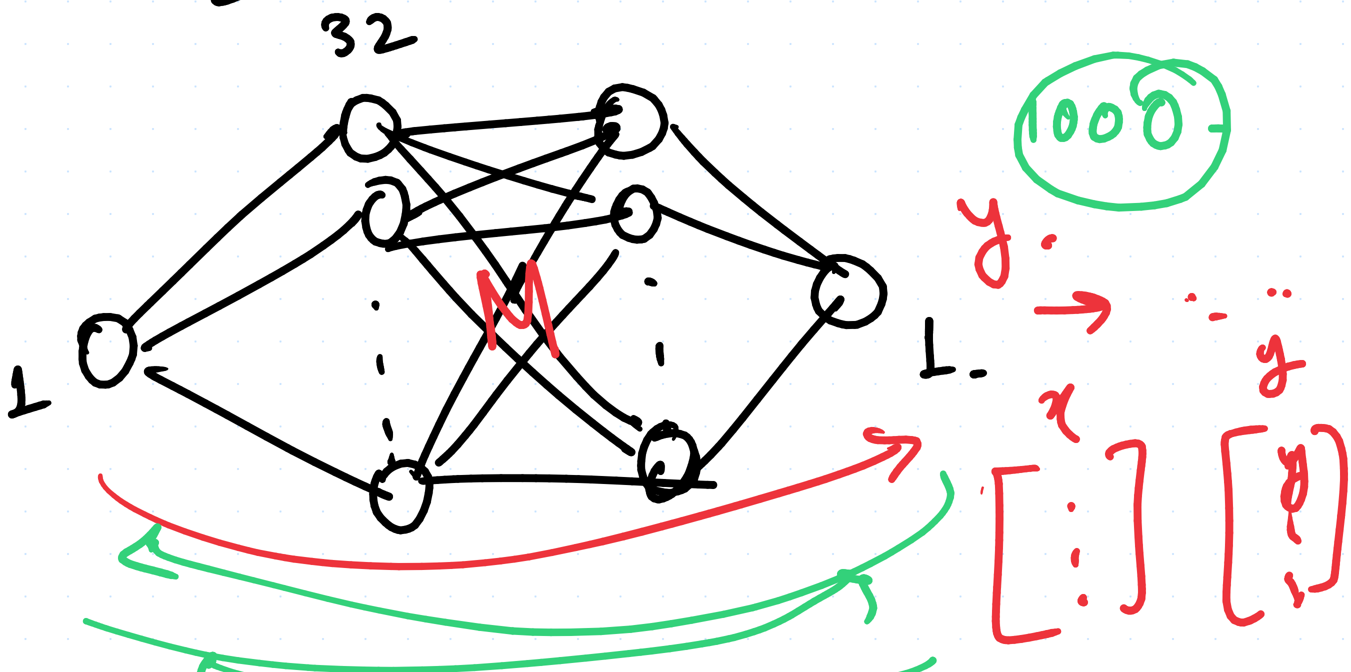
→ (0 – 1) values are always in b/w this.

→ $\sum x_i$ in softmax = 1

Simplex.







$\boxed{M}$

↓

CPU → GPU.

net.to(device).

net.train().

$y = net(x)$

$loss = MSE(\hat{y}, y).$