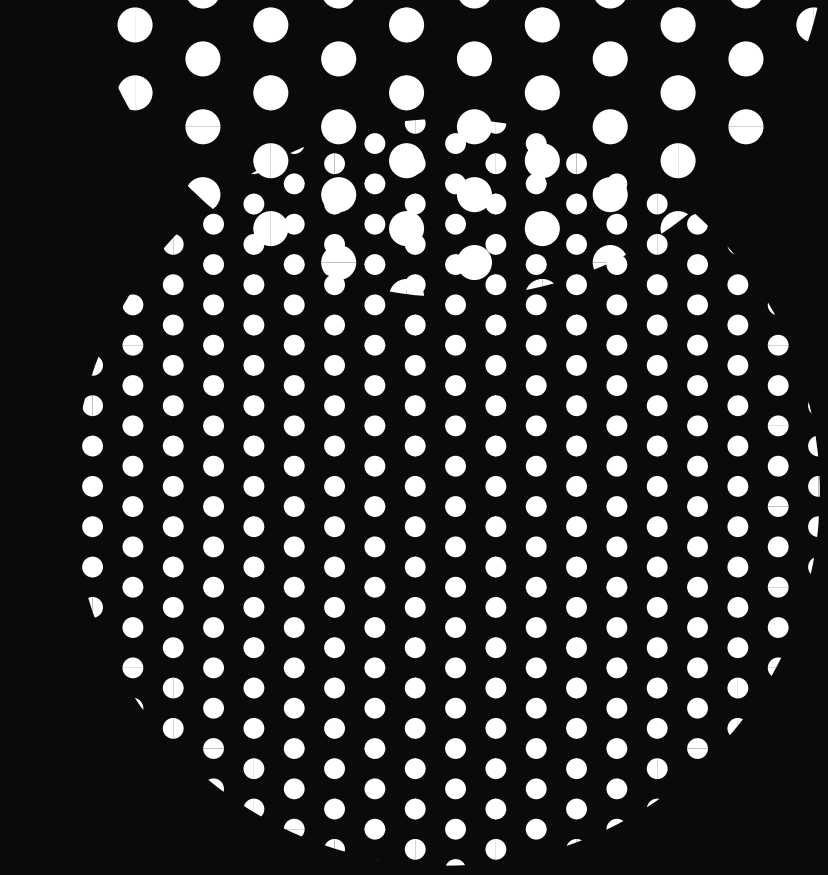



PyTorch

About Myself

SABBIR MOLLAH

- JAN 2020 – DECEMBER 2020

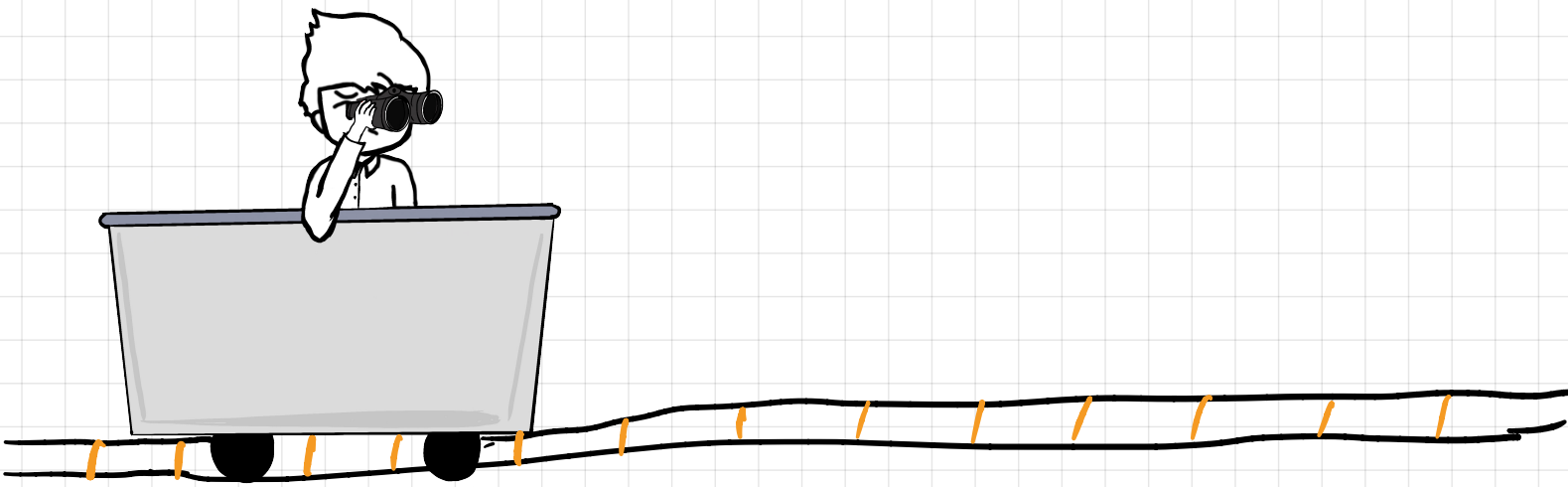
  Chair ⟶ NSU ACM SC

- 2021 – Current

  Part-time Research Assistant → Apurba-NSU R&D Lab

- 2022 – Current

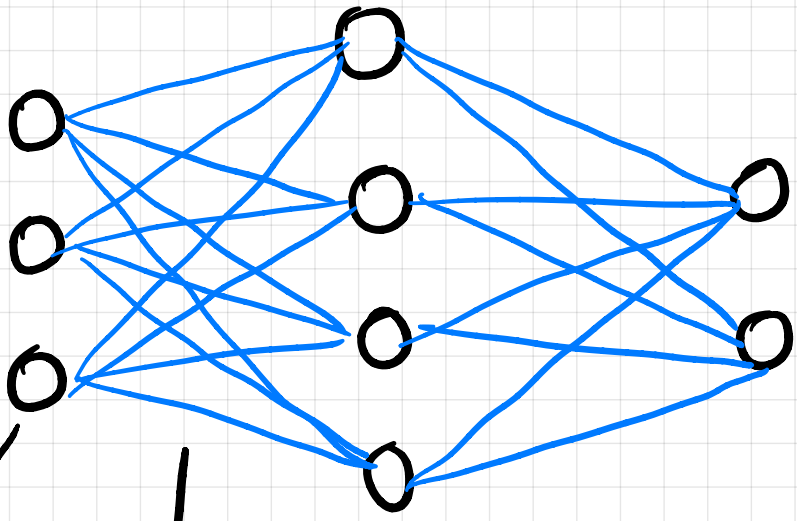
  LAB INSTRUCTOR → NSU ECE DEPT.



① Bias

② Activation

Why Bias?

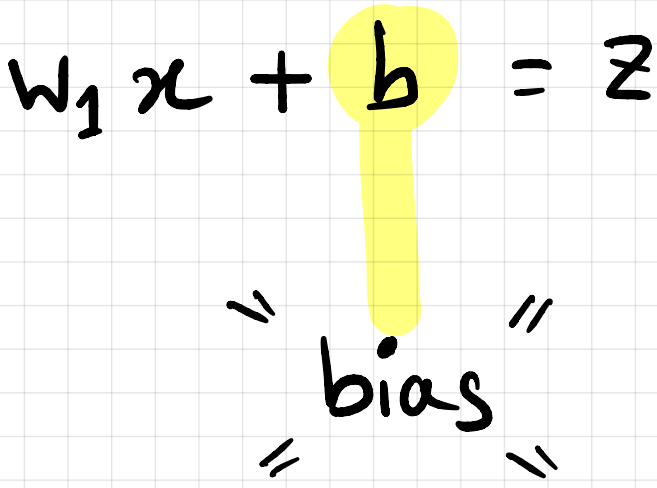


x W_1 z ...

3x1 4x3 4x1

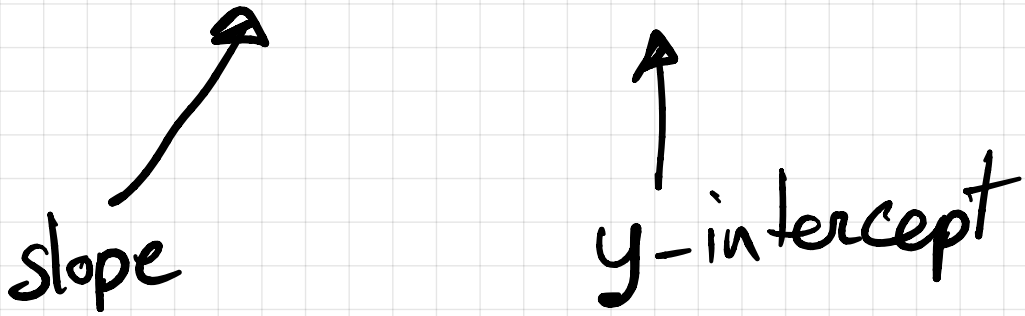
$$W_1 \cdot x = z$$

[4x3] [3x1] [4x1]


$w_1 x + b = z$
bias

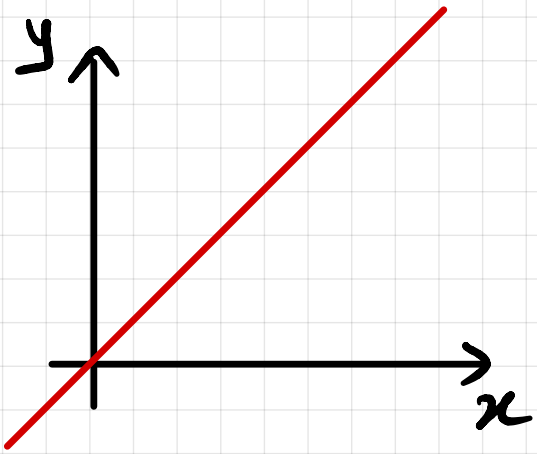
Equation of a line

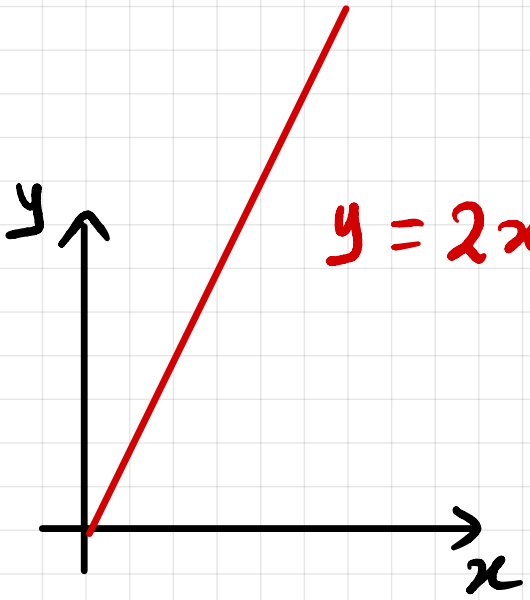
$$y = mx + c$$

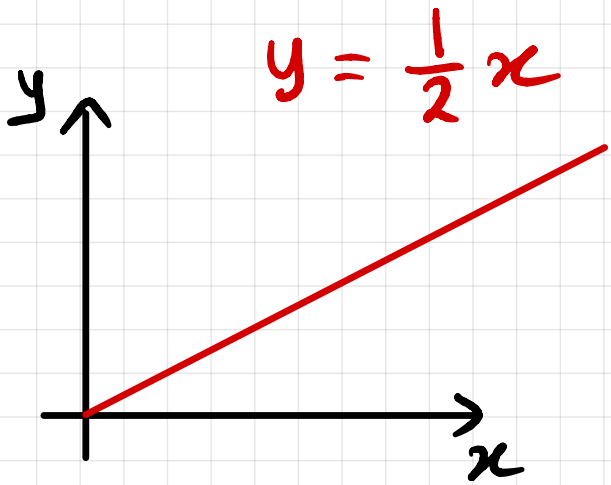


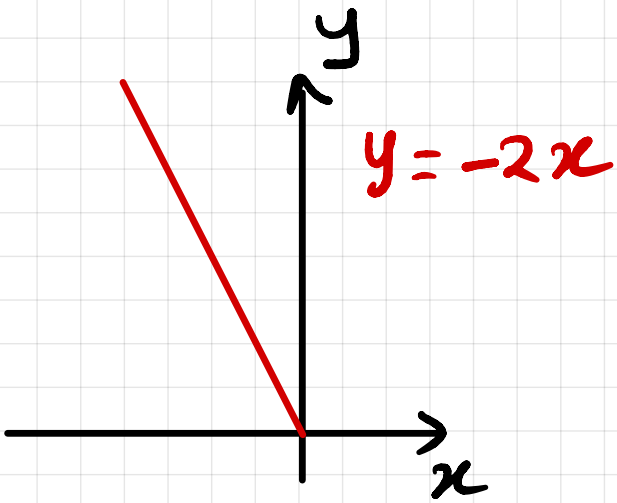
$y = mx$

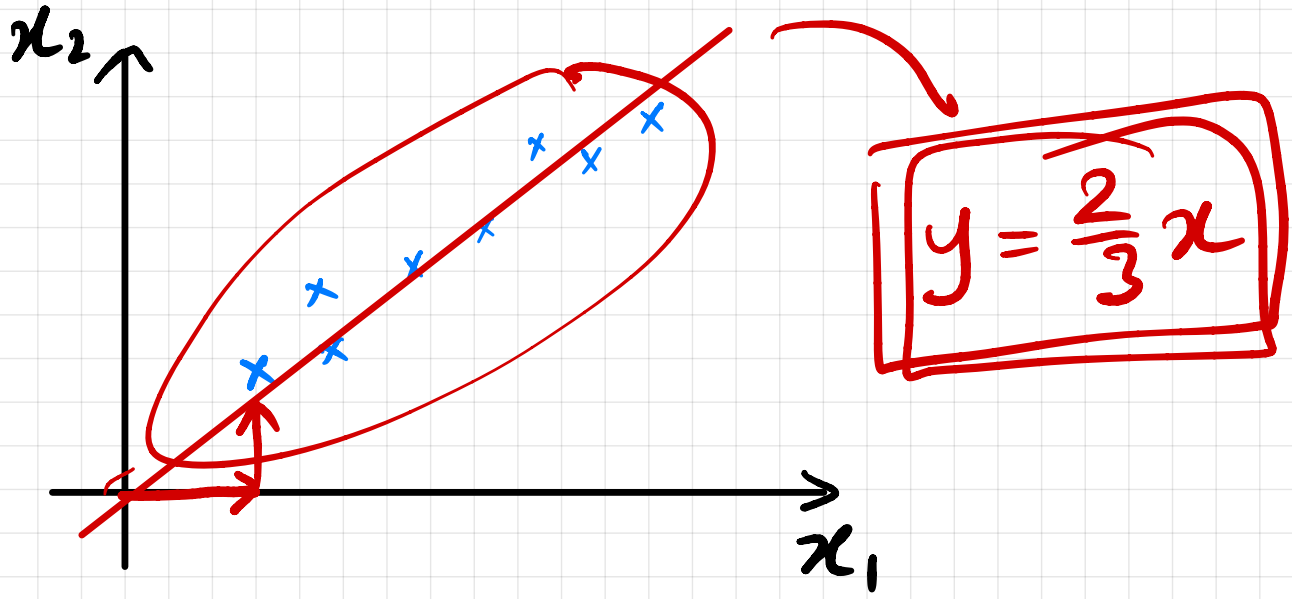
$y = 1 \cdot x$







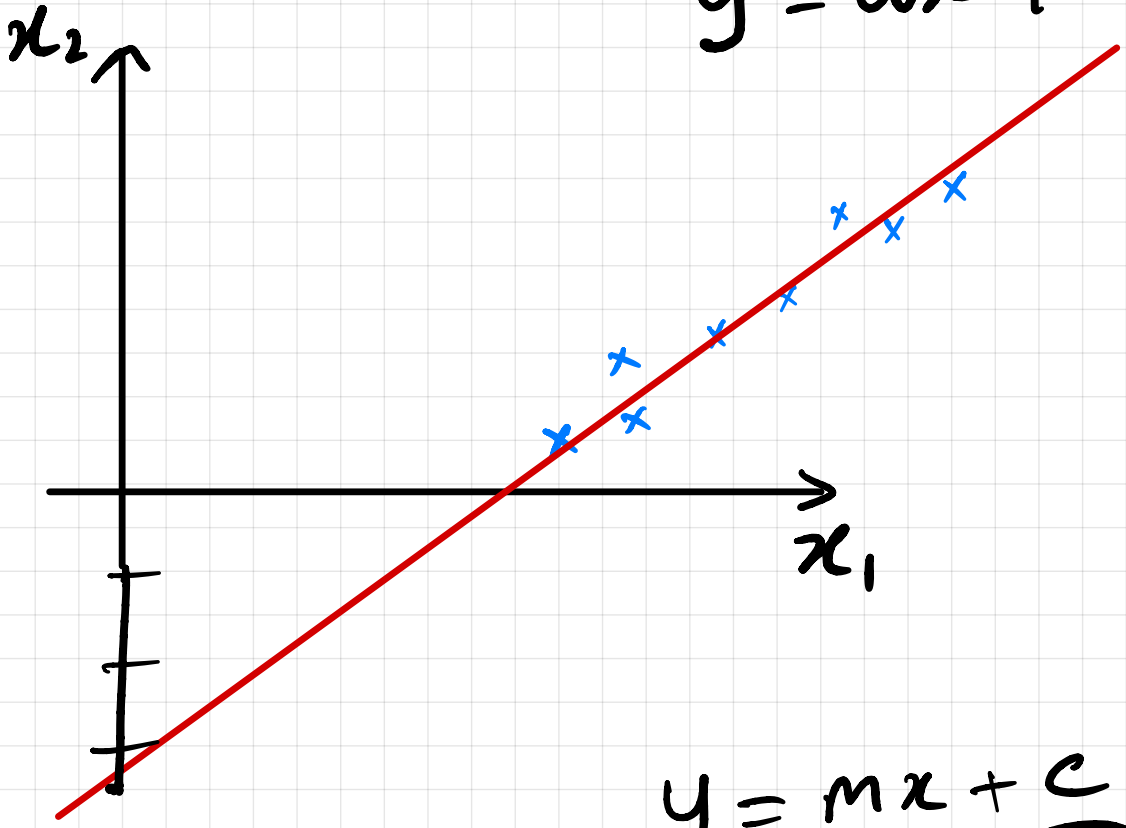




$$y = mx$$

bias

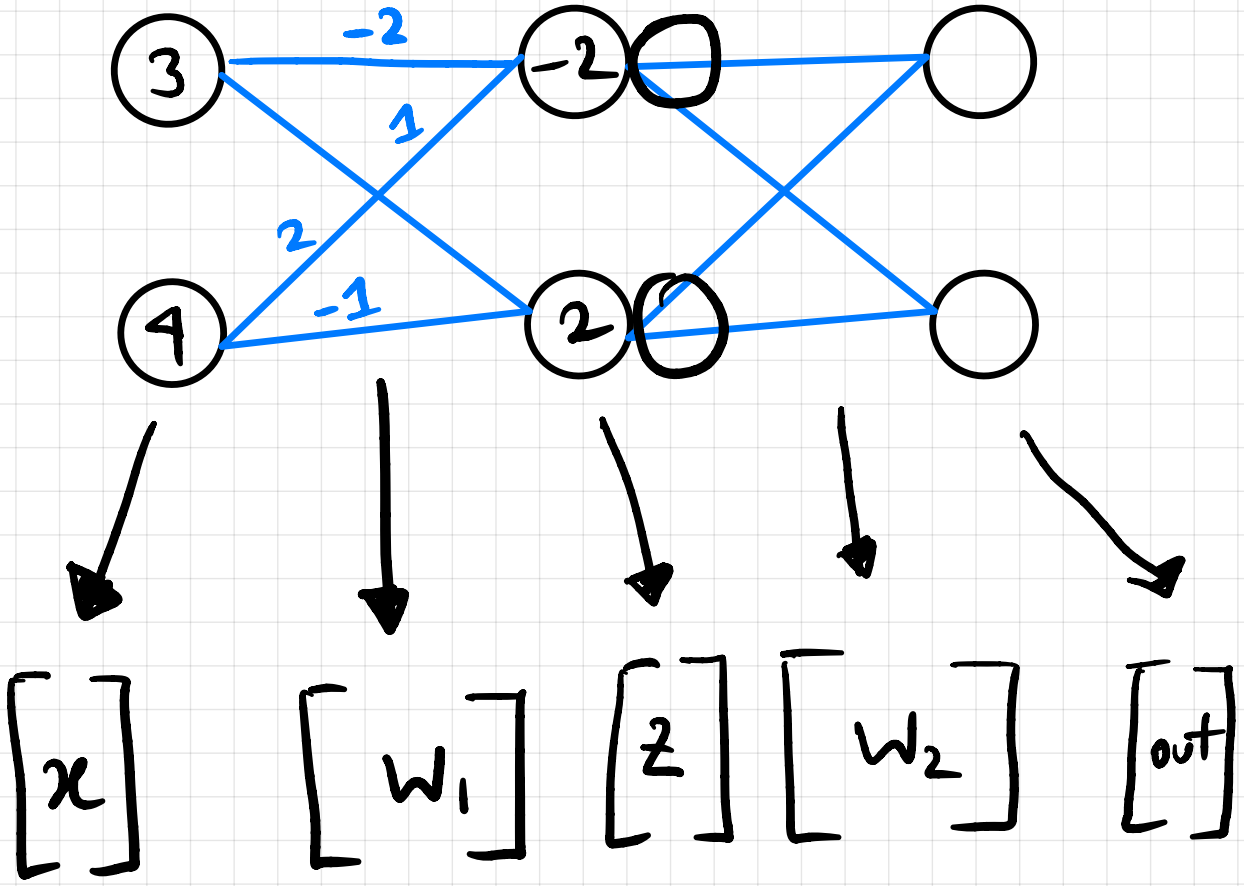
$$y = wx + b$$



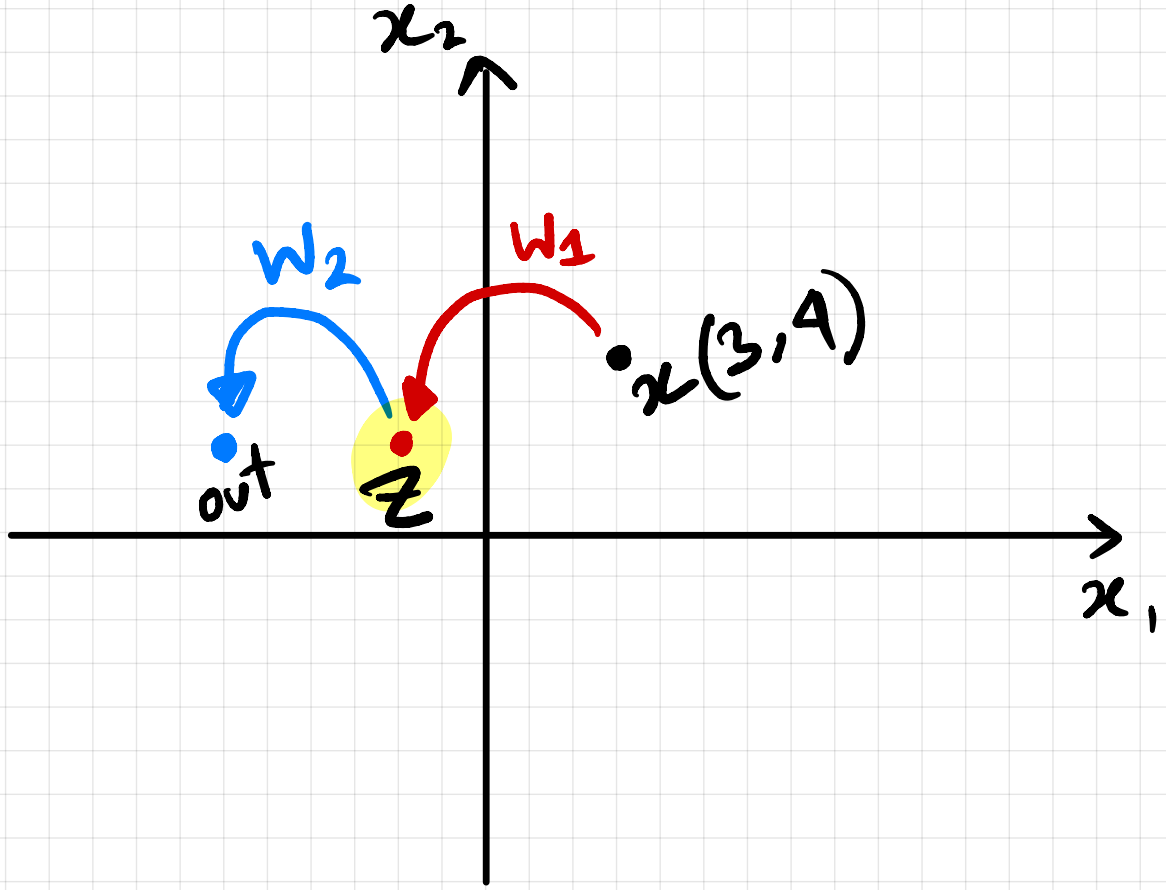
$$y = mx + c$$

$$y = \frac{2}{3}x - 7$$

Why activate ?



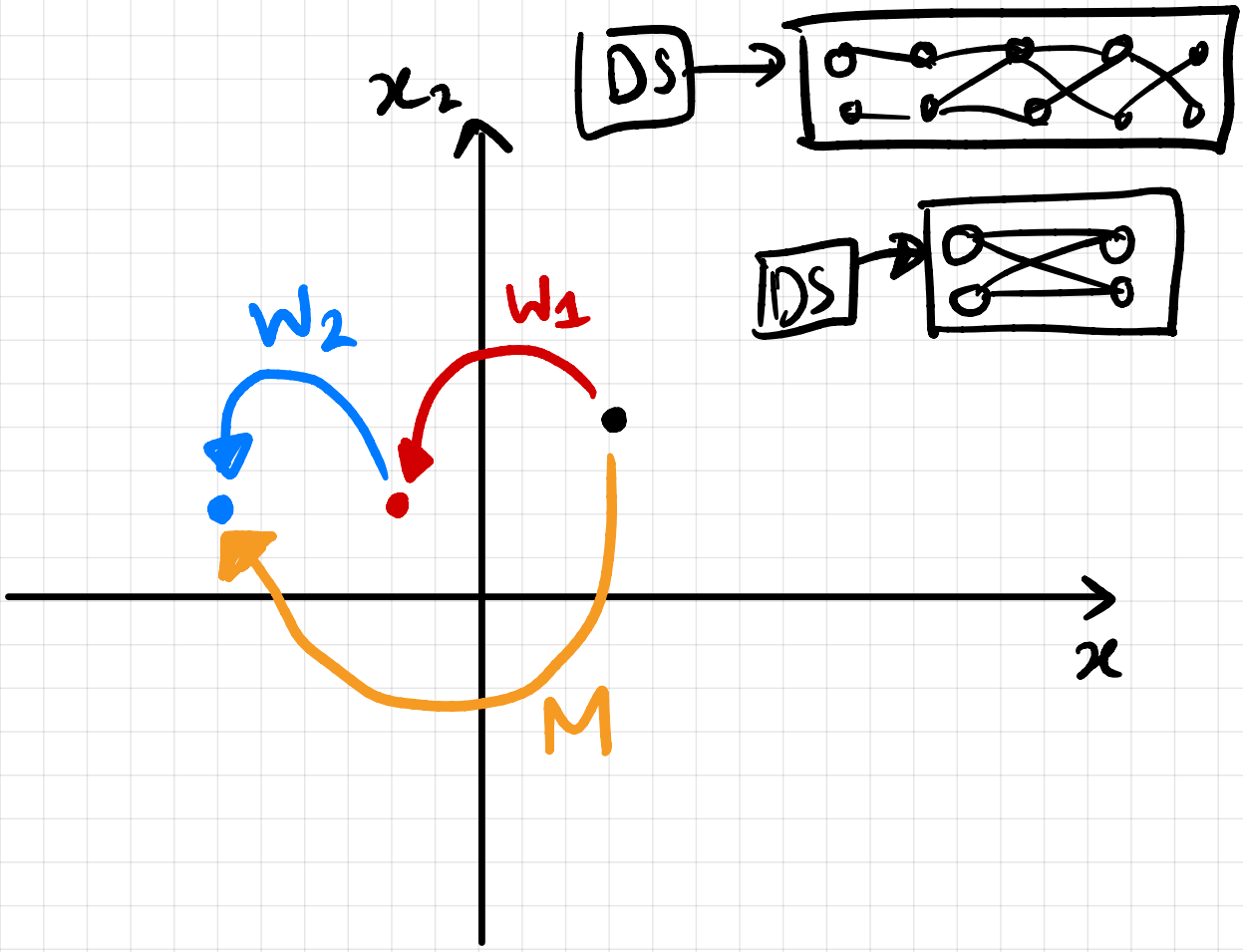
~~bias~~



$$z = W_1 \cdot x = \begin{bmatrix} -2 & 1 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} -2 \\ 2 \end{bmatrix}$$

$$out = W_2 * z = \begin{bmatrix} 1 & -2 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} -2 \\ 2 \end{bmatrix} = \begin{bmatrix} -6 \\ 2 \end{bmatrix}$$

$*$ ↑ matmul

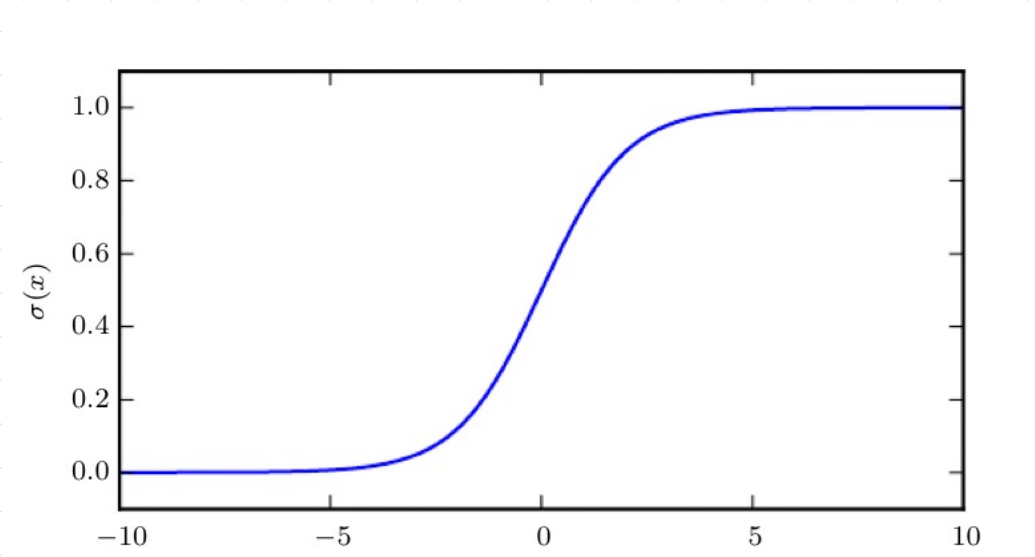


$$Mx = out$$

$$\begin{bmatrix} -2 & 0 \\ -6 & 5 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} -6 \\ 2 \end{bmatrix}$$

σ - Sigmoid Function

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



Properties:

1. Non Linear
2. [0-1] range
3. Differential

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

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

$$= -e^{-x}\left(-(1+e^{-x})^{-2}\right)$$

$$= e^{-x}(1+e^{-x})^{-2}$$

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

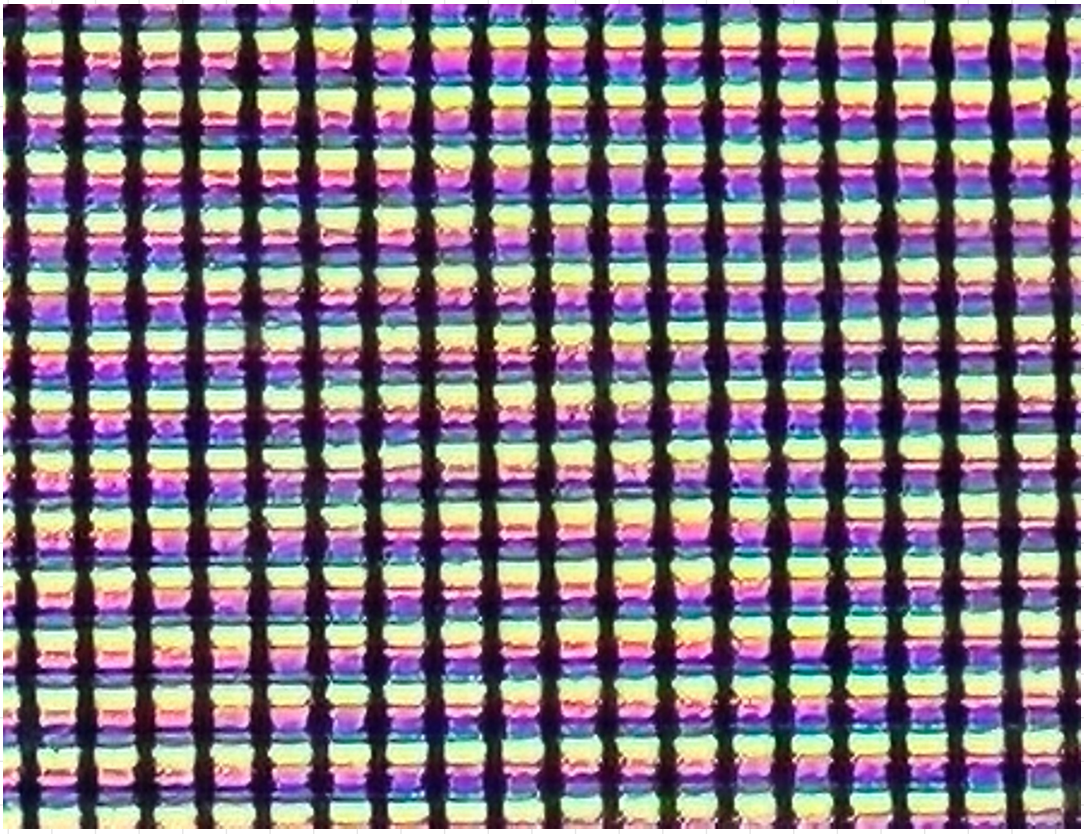
$$= \frac{1}{1+e^{-x}} \cdot \frac{e^{-x}}{1+e^{-x}}$$

$$= \frac{1}{1+e^{-x}} \cdot \frac{e^{-x}+1-1}{1+e^{-x}}$$

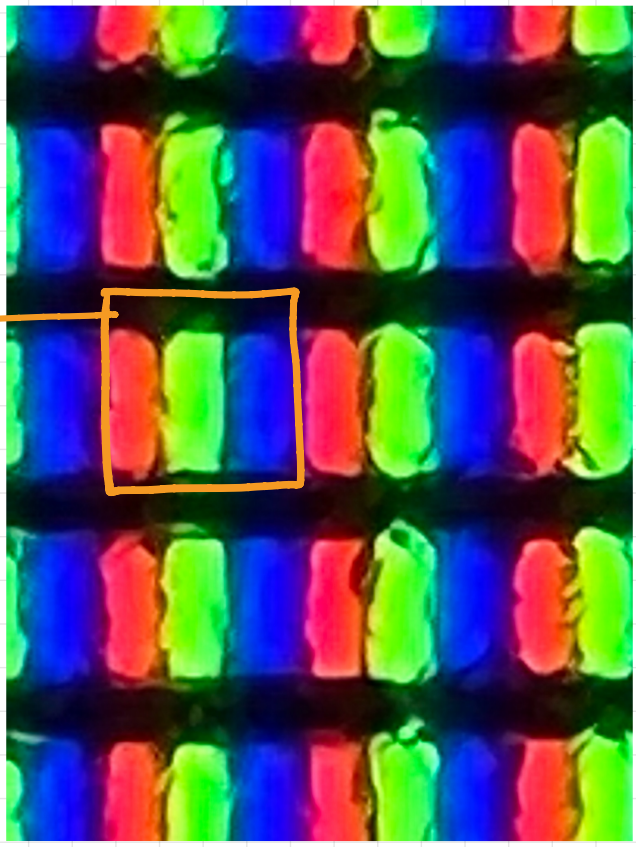
$$= \frac{1}{1+e^{-x}} \cdot \frac{e^{-x}+1}{1+e^{-x}} - \frac{1}{1+e^{-x}}$$

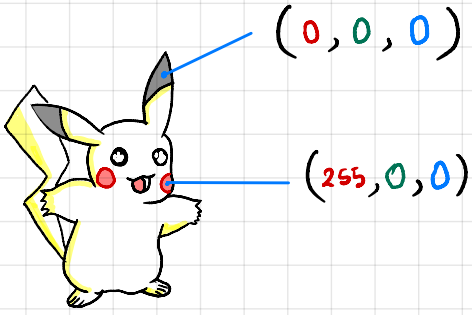
$$= \frac{1}{1+e^{-x}} \cdot \left(1 - \frac{1}{1+e^{-x}}\right)$$

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



Macro Camera
x 15 :





Tensors

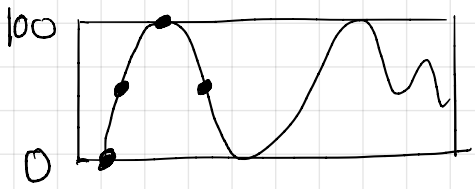
Dimension 0

Scalar Value

height = 75.5

Order 1

Mono Audio



[0, 50, 100, 50 ...]

Order 2

$$\begin{bmatrix} 0 & 255 \\ 255 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 100 \\ 255 & 255 \end{bmatrix}$$

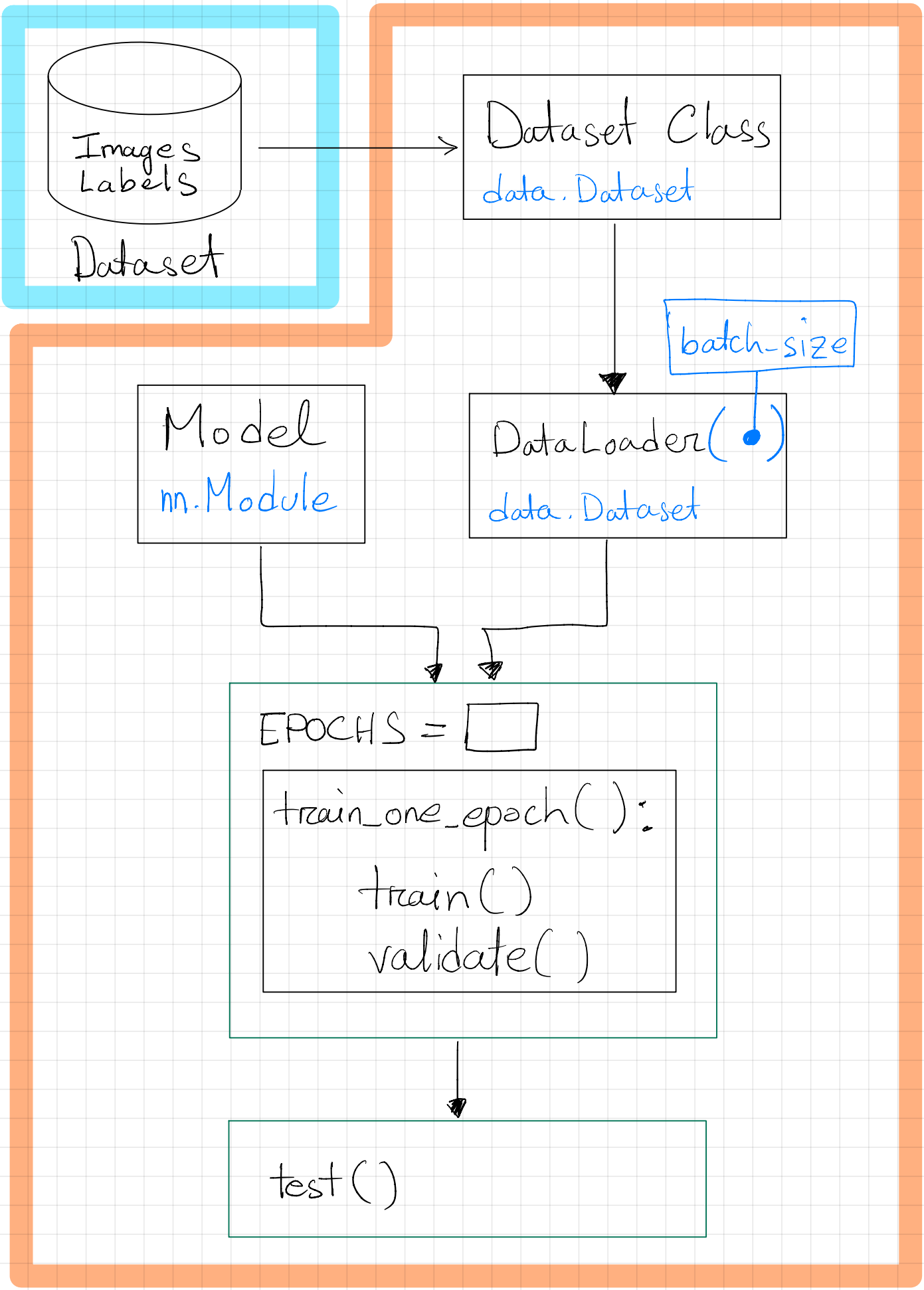
Order 3

$$\begin{bmatrix} 255 & 255 \\ 255 & 0 \end{bmatrix} \quad \begin{bmatrix} 0 & 255 \\ 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 0 & 255 \\ 255 & 255 \end{bmatrix}$$

R G B

RGB

PyTorch way :


Images
Labels
Dataset
Dataset Class
data.Dataset
batch-size
Model
nn.Module
DataLoader( )
data.Dataset
EPOCHS =
train_one_epoch( ):
train( )
validate( )
test( )

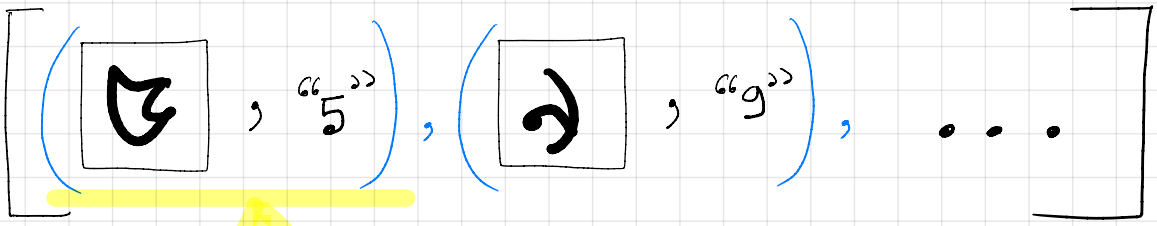
NUMTA - dataset

19702
LABELLED DATA

	filename	digit
0	a00000.png	5
1	a00001.png	3
2	a00002.png	1
3	a00003.png	7
4	a00004.png	0

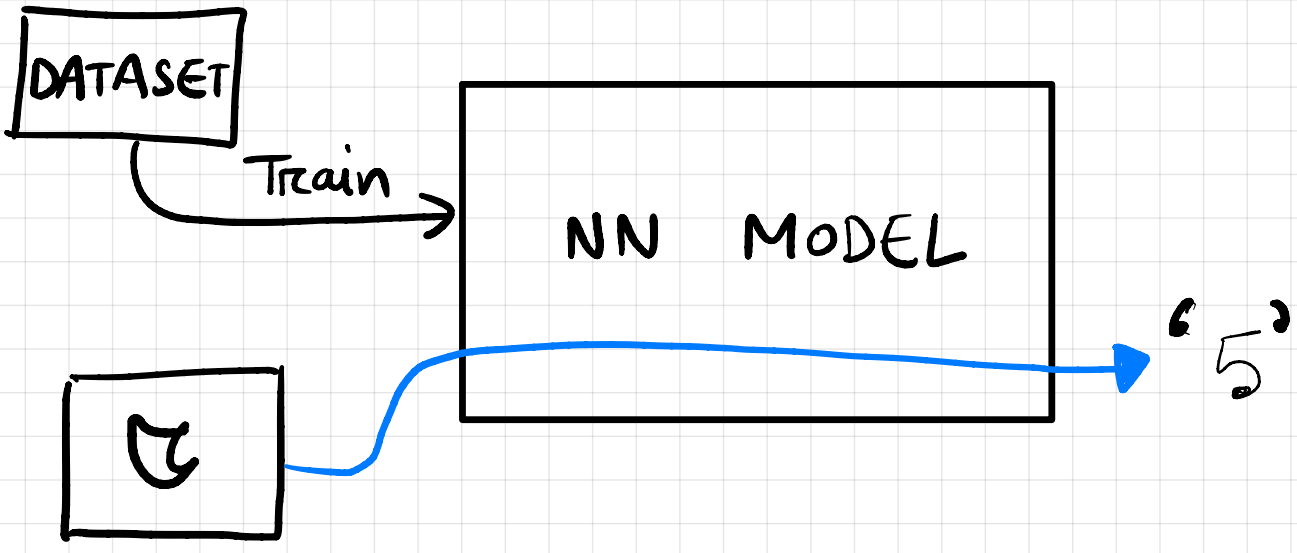
CSV

dataset =

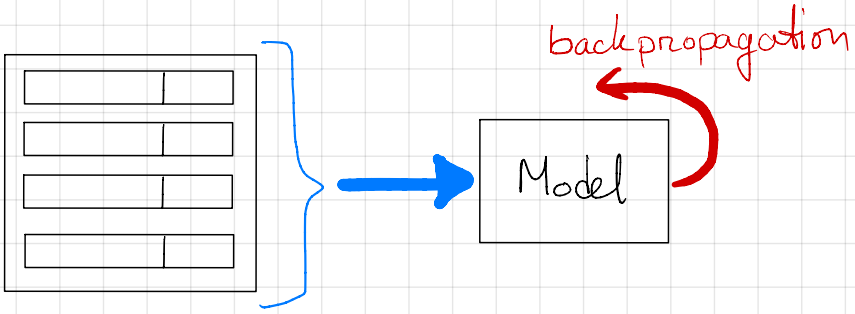


① dataset [0]

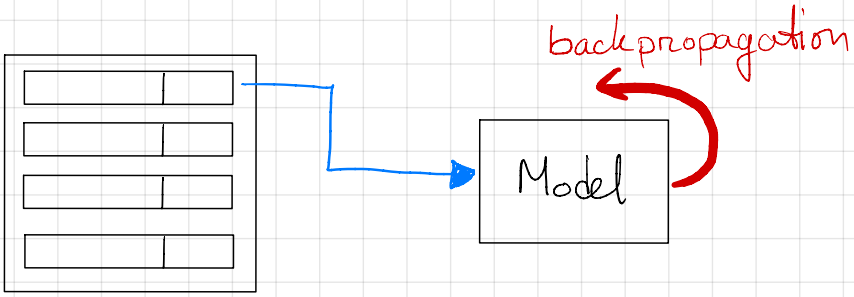
② len(dataset) → 60_000



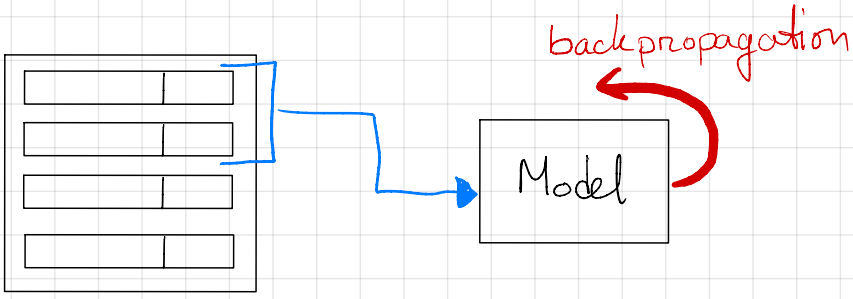
(1) Batch Gradient Descent



(2) Stochastic Gradient Descent

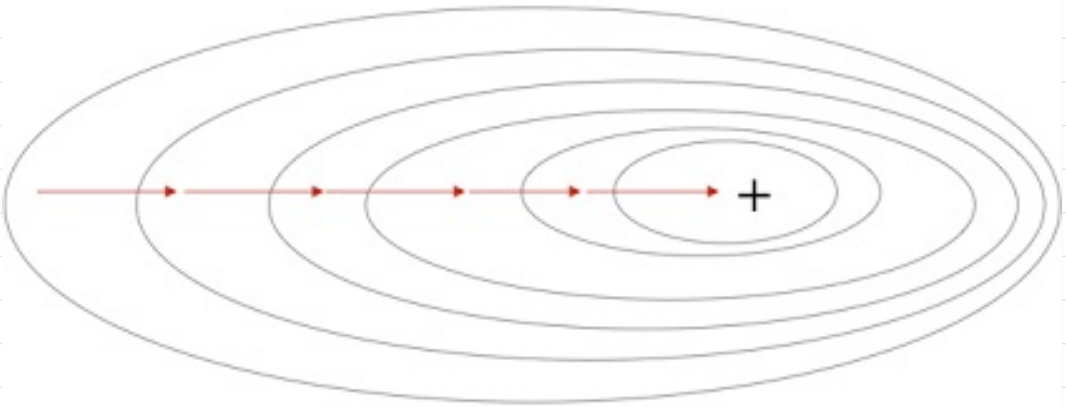


(3) Mini Batch Gradient Descent

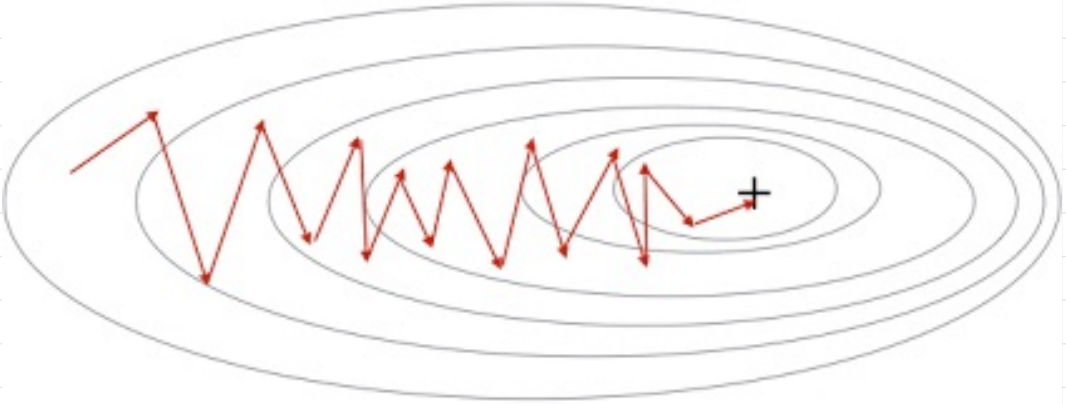


batch_size → 2

Gradient Descent



Stochastic Gradient Descent



Mini-Batch Gradient Descent