

Weight initialization

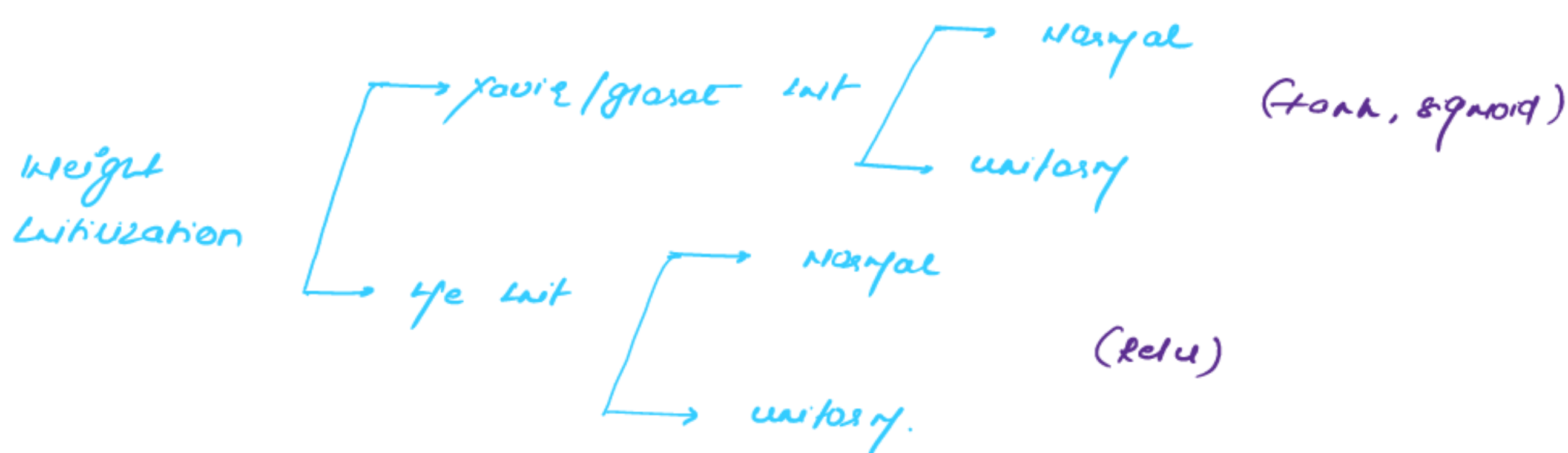
30 October 2023

07:06

model weight assign. randomly.

things we need to avoid.

- ① weights should not be initialize to zero. (NO Training)
- ② Non zero initialization. (Linear model).
- ③ Random initialization. small weights. (vanishing gradient)
- ④ Random initialization large weights. (exploding gradient)



$$\text{Variance} = \frac{1}{n}$$

[sq]

✓ $n \uparrow$ $\sigma \downarrow$

$n \downarrow$ $\sigma \uparrow$



{ fan in, fan out }

Xavier's Init
(Normal)

$$\sqrt{\frac{1}{\text{fan in}}}$$

$$\sqrt{\frac{2}{\text{fan in} + \text{fan out}}}$$

He Init
Normal

$$\sqrt{\frac{2}{\text{fan in}}}$$

Xavier uniform
uniform

$$\sqrt{\frac{6}{\text{fan in} + \text{fan out}}}$$

He uniform

$$\sqrt{\frac{6}{\text{fan in}}}$$

top

kernel_initializer = 'xxx'

= he-normal } ReLU
= he-uniform }

= glorot-normal } tanh
= glorot-uniform } sigmoid

By default technique used.