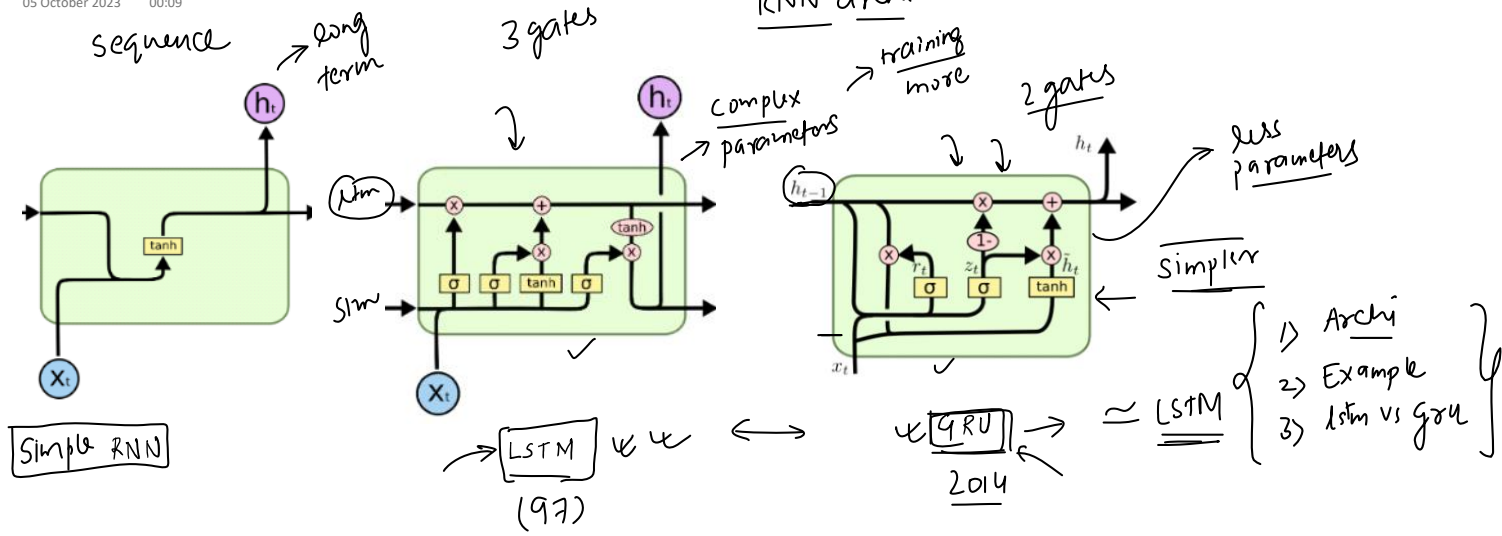


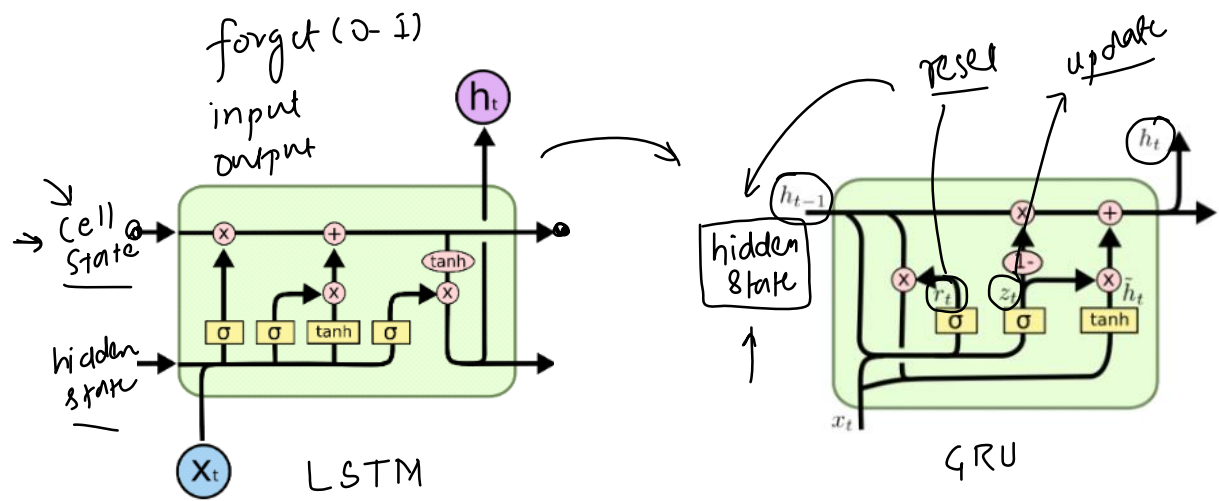
What is GRU

05 October 2023 00:09



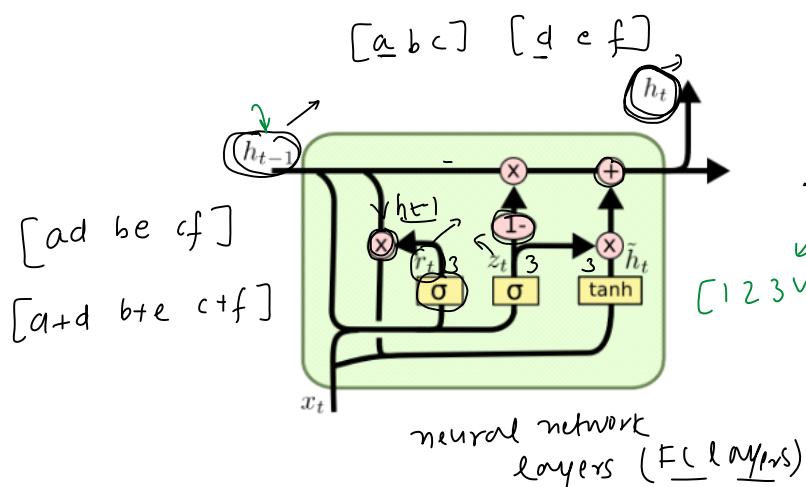
The Big Idea Behind GRU

05 October 2023 00:47



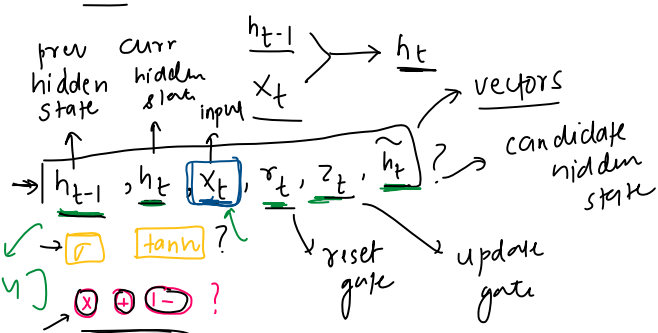
The Setup

05 October 2023 01:07



→ Advise → LSTM/GRU → confusing

goal → $[t]$



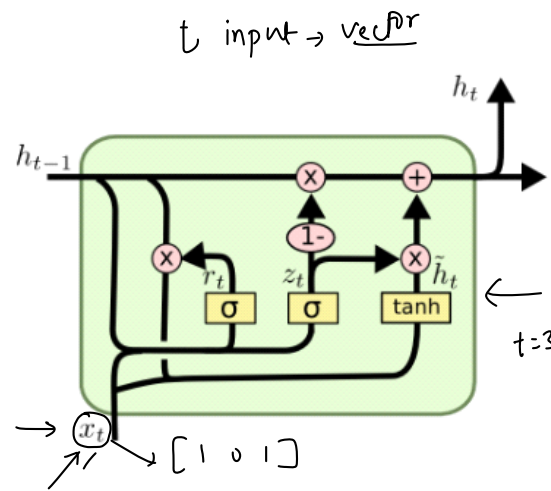
$[1 \ 2 \ 3 \ 4]$ 4dim same

number of
= 5 6

$\sigma \rightarrow \sigma \ \sigma \ \sigma \ \sigma \ \sigma$

The Input x_t

05 October 2023 01:52



Sentiment

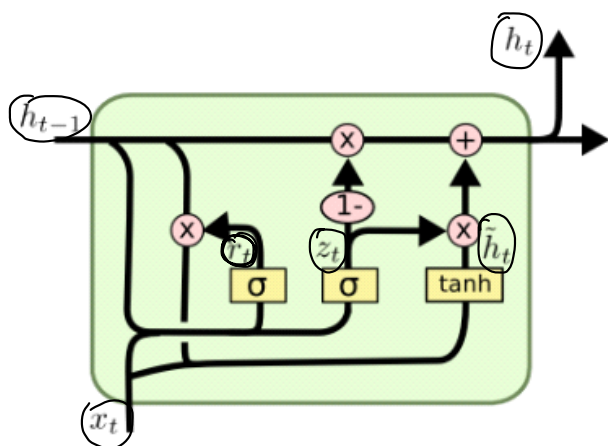
ONE / BOW / Wordvec

text	review
cat mat rat	1
cat rat rat	0
mat rat mat	1

S1	[1 0 0]	[0 1 0]	[0 0 1]
S2	[1 0 0]	[0 0 1]	[0 0 1]
S3	[0 1 0]	[0 0 1]	[0 1 0]

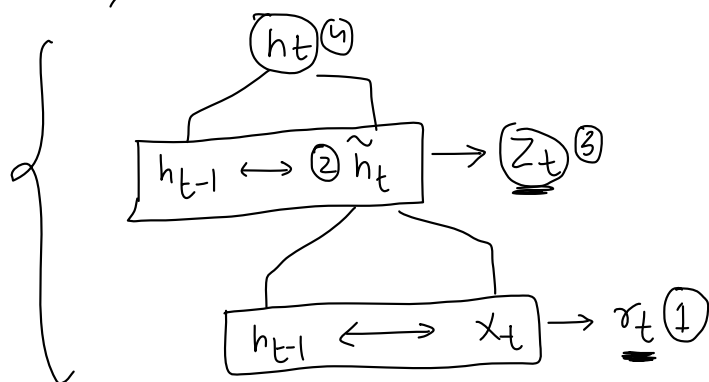
Architecture

05 October 2023 02:10



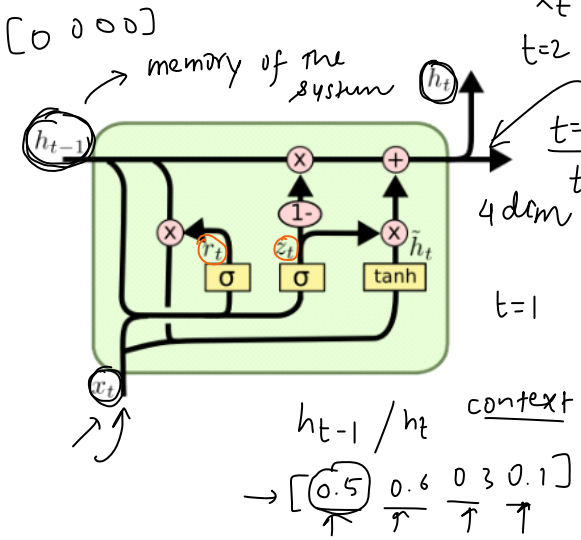
$$h_{t-1}, x_t \rightarrow h_t$$

- 1) Calculate $\underline{r_t}$ (reset gate)
- 2) Calculate \tilde{h}_t (candidate hidden state)
- 3) Calculate $\underline{z_t}$ (update gate)
- 4) Calculate $\underline{h_t}$ (current hidden state)



What exactly is hidden state?

05 October 2023 02:19



- memory \rightarrow context
- h_{t-1}
- happy/sad
- There was a king Vikram very strong and powerful
 - There was an enemy king kaali
 - Both had a war and kaali killed Vikram
 - Vikram had a son Vikram Jr who grew up he to become very strong just like his father
 - He also attacked Kaali But got killed
 - Vikram Jr too had a son called Vikram super Jr and when he grew up he also fought kaali
 - And he killed kaali and took revenge of his father and grand father

h_{t-1}

\rightarrow [Power, Conflict, Tragedy, Revenge]

$\rightarrow [0.9 \ 0.6 \ 0.9 \ 0.9]$

$\rightarrow [0.9 \ 0 \ 0 \ 0] h_t$

$\rightarrow [1 \ 0.5 \ 0 \ 0] \rightarrow h_{t-1}$

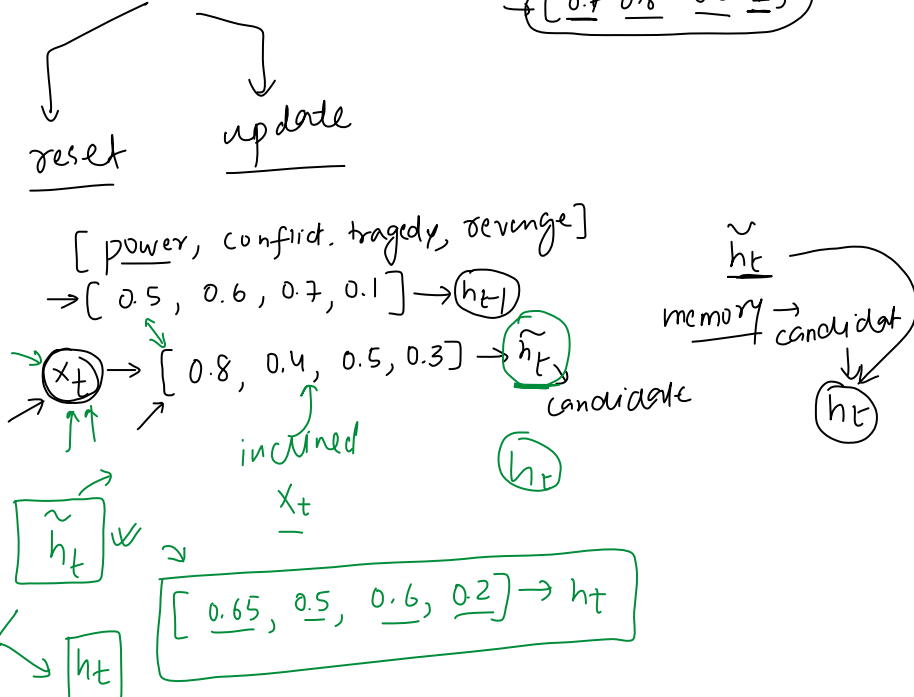
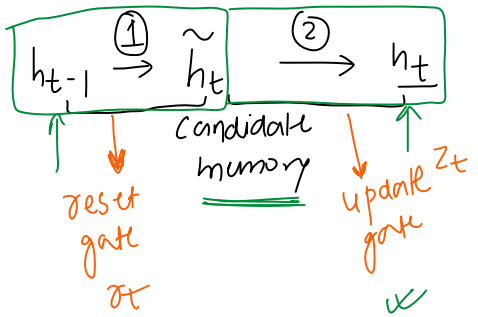
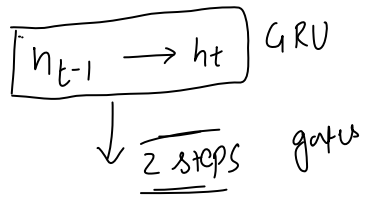
$\rightarrow [0.7 \ 0.7 \ 0.4 \ 0]$

$\rightarrow [0.9 \ 0.6 \ 0.3 \ 0]$

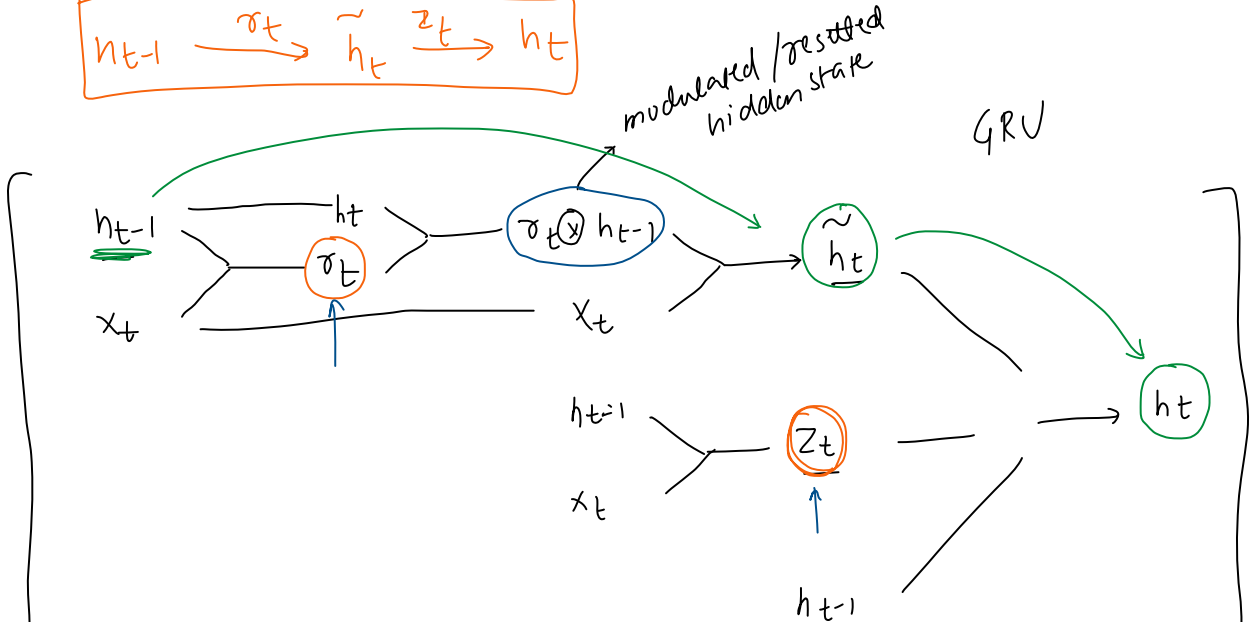
$\rightarrow [0.8 \ 0.8 \ 0.6 \ 0.5]$

$\rightarrow [1 \ 0.9 \ 0.5 \ 0.7]$

$\rightarrow [0.7 \ 0.8 \ 0.3 \ 1]$



$h_{t-1} \xrightarrow{\sigma_t} \tilde{h}_t \xrightarrow{z_t} h_t$



|

h_{t-1} /

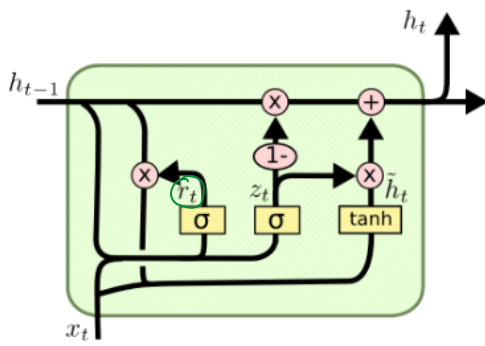
]

Calculating the reset gate

05 October 2023 14:24

(weighted sum)

0 1

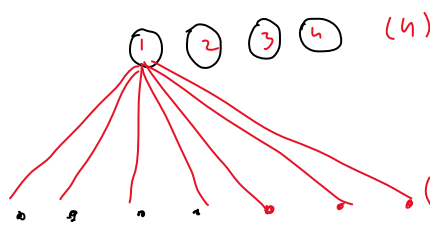
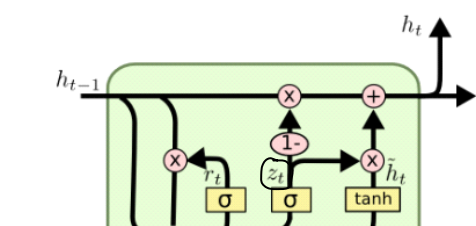
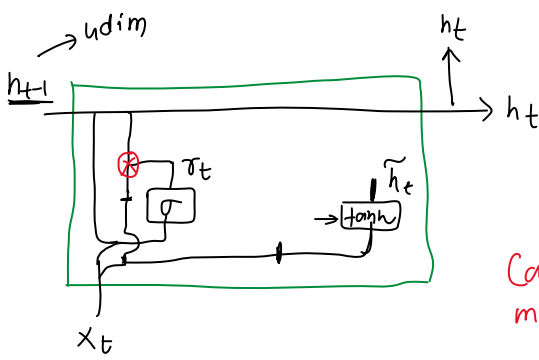
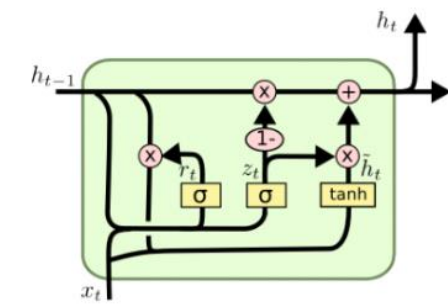
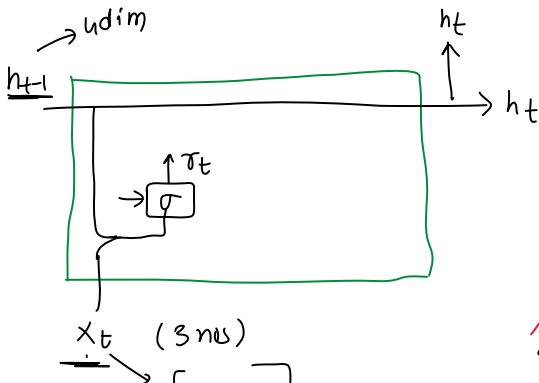


reset

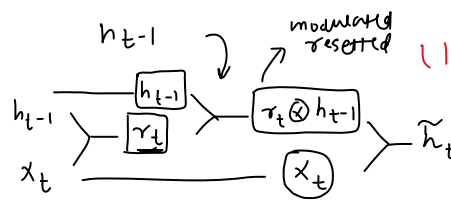
$r_t \rightarrow$ vector $\rightarrow h_{t-1}$
gate

$h_{t-1} = [0.6, 0.6, 0.7, 0.1]$
 $r_t = [0.8, 0.2, 0.1, 0.9]$
gate
50% percent 20% 90%

$$r_t = \sigma(W_r[h_{t-1}, x_t] + b_r)$$



$$(7 \times 4) = 28 + 4$$

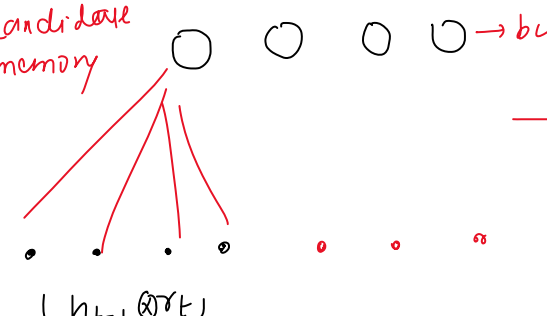


$h_{t-1} = [0.6, 0.6, 0.7, 0.1]$
 $r_t = [0.8, 0.2, 0.1, 0.9]$
power 90% confider 70% tragary 70%

$$h_{t-1} \otimes r_t = [0.48, 0.12, 0.07, 0.09]$$

$$\tilde{h}_t = \tanh(W_c[h_{t-1} \otimes r_t, x_t] + b_c)$$

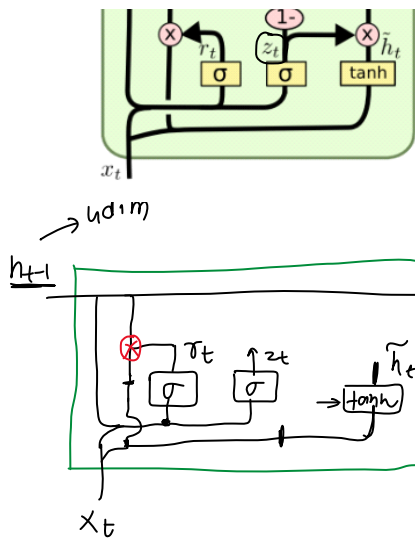
Candidate memory



$$W_c (7 \times 4)$$

$$\tilde{h}_t = [0.7, 0.2, 0.1, 0.2]$$

$$h_{t-1} = [0.6, 0.6, 0.7, 0.1]$$



$$z_t = \sigma(W_z[h_{t-1}, x_t] + b_z)$$

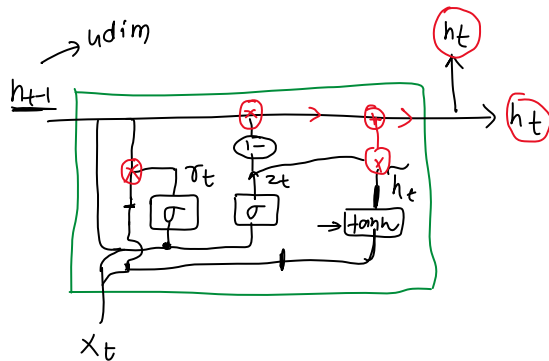
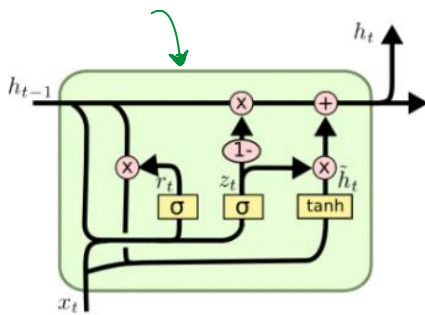
gate

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t$$

$x_t \rightarrow \text{important}$
 $x_t \rightarrow \text{not imp}$

(1×4)
 (7×4)

$\rightarrow \text{bias}(4)$
 $\rightarrow (7 \times 4)$



$$\tilde{h}_t = [0.7 \ 0.2 \ 0.1 \ 0.2]$$

$$h_{t-1} = [0.6 \ 0.6 \ 0.7 \ 0.1]$$

$$z_t = [0.1 \ 0.7 \ 0.8 \ 0.2]$$

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t$$

$(1 - z_t) \rightarrow 0.9$
 $[0.9 \ 0.3 \ 0.2 \ 0.8] [0.6 \ 0.6 \ 0.7 \ 0.1]$
 $[0.54 \ 0.18 \ 0.14 \ 0.08]$
 $\Rightarrow h_t = [0.61 \ 0.32 \ 0.22 \ 0.12]$
 power output fragility runway

$$h_{t-1} \text{ and } x_t \rightarrow z_t \odot h_{t-1} \rightarrow \tilde{h}_t$$

$$h_{t-1} \text{ and } x_t \rightarrow z_t$$

$$h_{t-1} \text{ and } z_t \rightarrow h_t$$