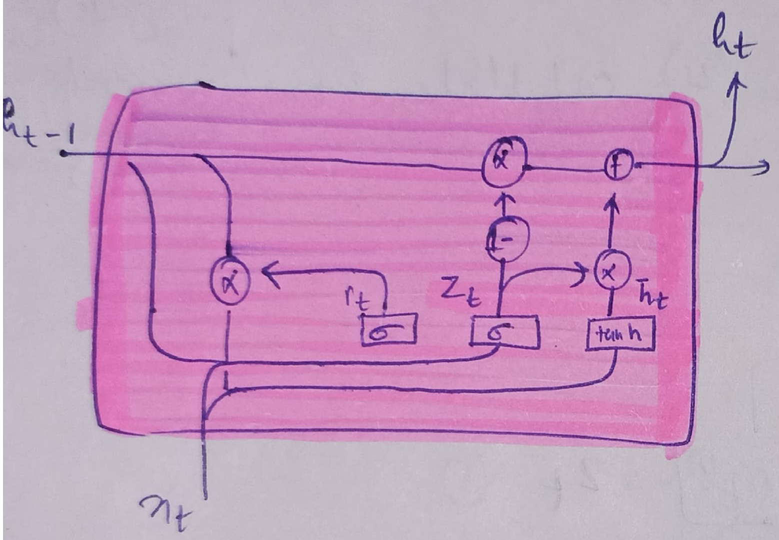
* GRU → Gated Recurrent Unit :

- less Training Time.
- less parameters.
- simple Arcitecture
- Two gates as compared to lstm.



→ The input x_t

Input → vector

Sentiment	
text	vector
cat mat rat	1
cat rat mat	0
mat rat cat	1

ONE hot encoding

unique vocabulary → 3

cat mat rat

S_1 [1 0 0] [0 1 0] [0 0 1]

S_2 [1 0 0] [0 0 1] [0 1 0]

S_3 [0 1 0] [0 0 1] [1 0 0]

→ Advise → LSTM / GRU → Confusing

→ Goal → $\boxed{t}$

h_{t-1}, x_t → h_t

→ h_{t-1}, h_t, x_t, r_t, z_t,

Prev hidden state, current hidden, Time step, reset gate, update gate

candidate hidden state — $\tilde{h}_t$ → ?

[σ] [tanh] ?

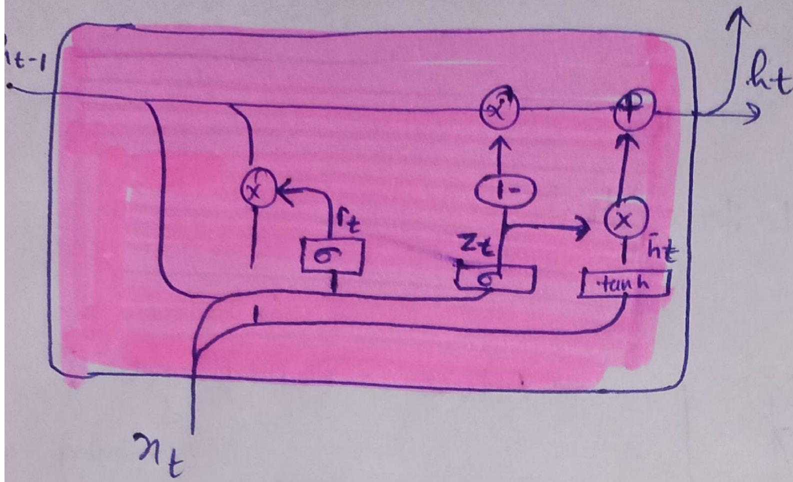
(×) (+) (1-) ?

→ dimension always same except x_t.

→ [σ] [tanh] → neural network layer.

→ (×) (+) (1-) → pointwise operators.

Architecture :

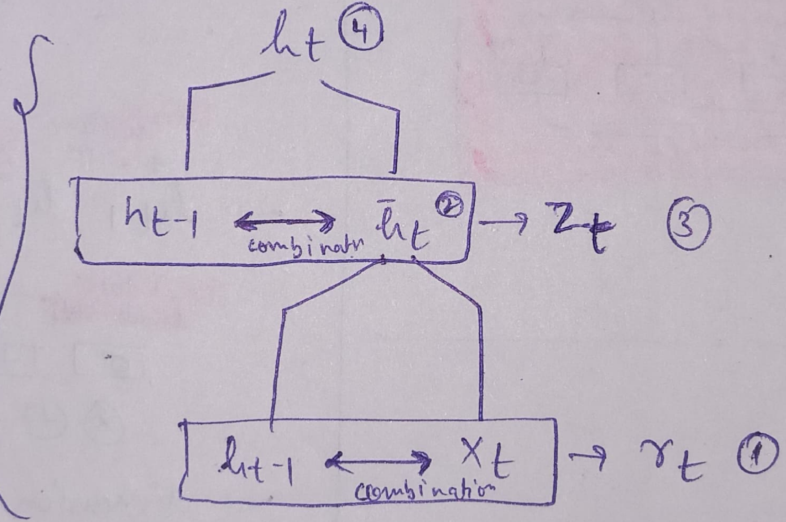


find out !!

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

Steps

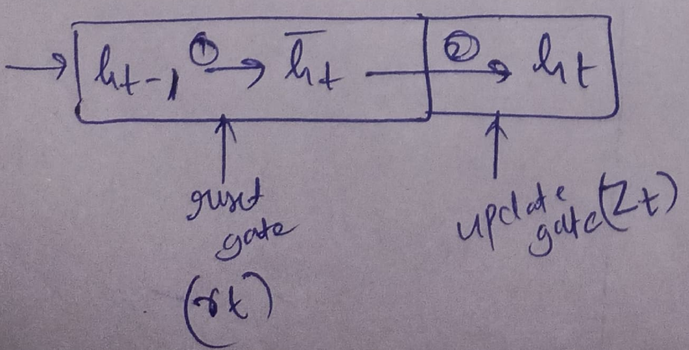
1) calculate r_t → (reset gate)
2) calculate $\bar{h}_t$ (candidate hidden state)
3) calculate z_t (update gate)
4) calculate h_t (current hidden gate).

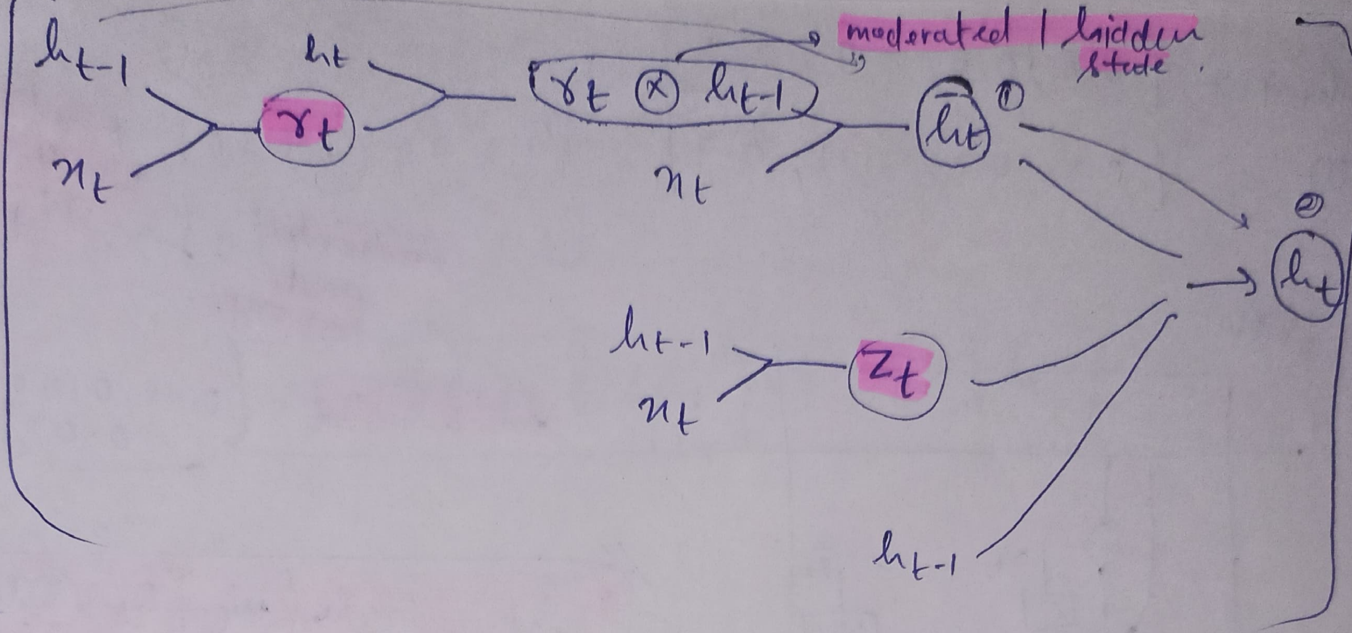


→ hidden state is Just like a memory of the system.
content (store).

→ $\underbrace{h_{t-1} \xrightarrow{1} \bar{h}_t}_{\text{candidate memory}} \xrightarrow{2} h_t$

$\bar{h}_t$
memory → candidate $\bar{h}_t$





* **Calculate the reset gate :-**

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

reset (lower)   reset (higher).

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

$r_t = [0.8, 0.2, 0.1, 0.9]$

80% percent   20%,   90%...

gate

①

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

[P power, C conflict, t tracking, r revenge]:

[- - - -]

r_t value $\rightarrow$ (0-1).

{Vikram, Kaali, war} h_{t-1}

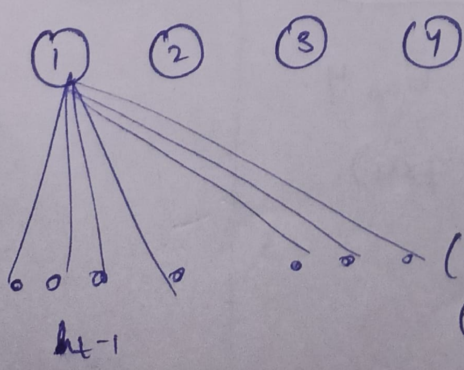
Vikram Jr x_t — dimension reset.

Matlab x_t क्या करेगा। h_{t-1} के dimension को Reset करेगा!!

r_t का काम Past memory से कुछ चीजों को Reset करना Based on current Input.

$h_{t-1} \rightarrow$ 4 dimension

$x_t \rightarrow$ 3 (no. of nodes)



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

$(7) \times 1 \rightarrow (1 \times 7)$

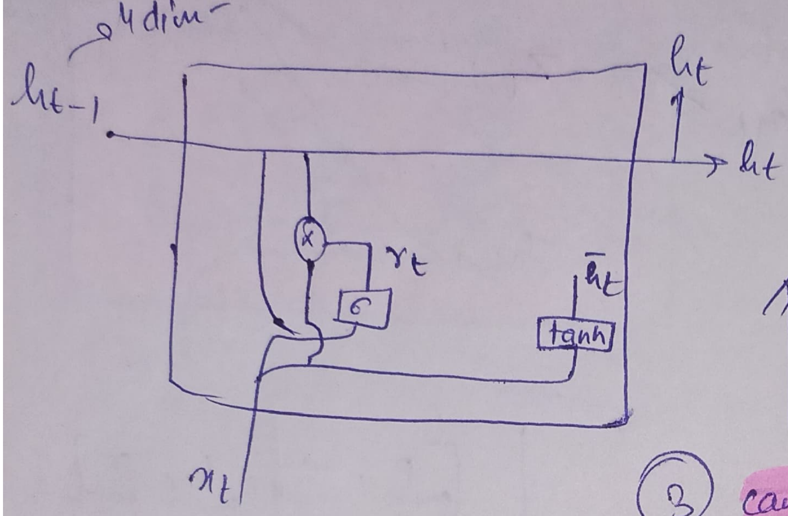
$(1 \times 7)\ (7 \times 4)$

$(1 \times 4) + (1 \times 4)$

$\rightarrow \sigma(1 \times 4)$

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

$r_t = [0.8, 0.2, 0.1, 0.9]$.



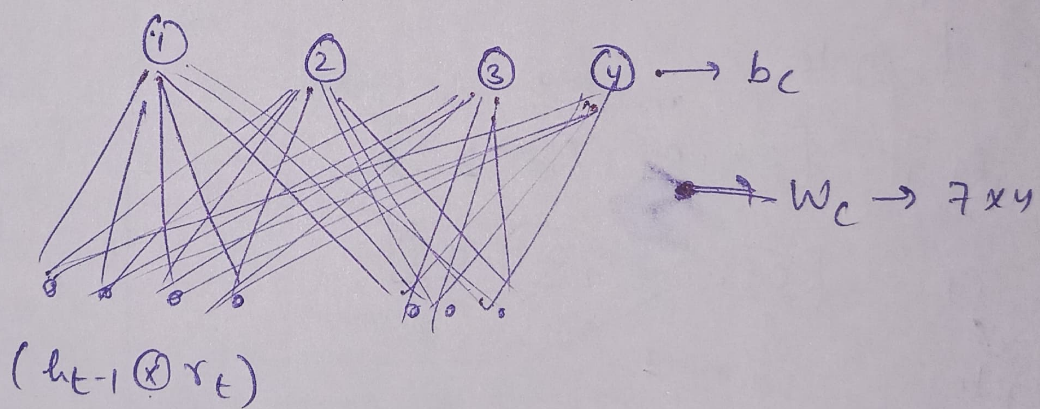
(2)

modulated past memory $\rightarrow$ based on current input x_t

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

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

(3) candidate memory



---

h_{t-1}, x_t $\rightarrow z_t \rightarrow$ update gate

$$\boxed{z_t \rightarrow \sigma\left(W_z\left(h_{t-1}, x_t\right) + b_z\right)}$$

○ ○ ○ ○ $\rightarrow$ bias 4.

$\rightarrow (7 \times 4)$

∘ ∘ ∘ ∘ ∘ ∘ ∘

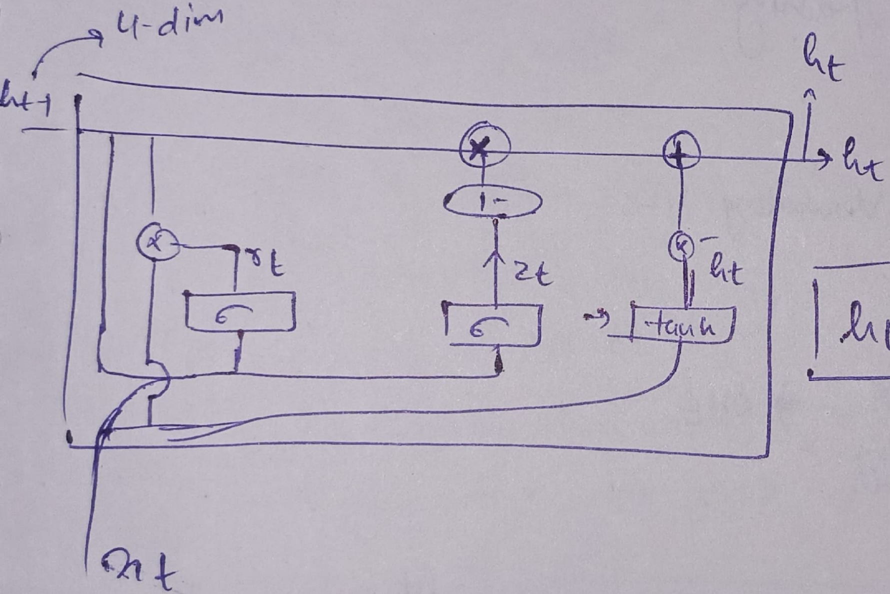
h_{t-1}

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

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

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

$$\boxed{h_t = (1 - z_t) \otimes h_{t-1} \oplus z_t \otimes \bar{h}_t}$$



$$\boxed{h_t = (1 - z_t) \otimes h_{t-1} \oplus z_t \otimes \bar{h}_t}$$