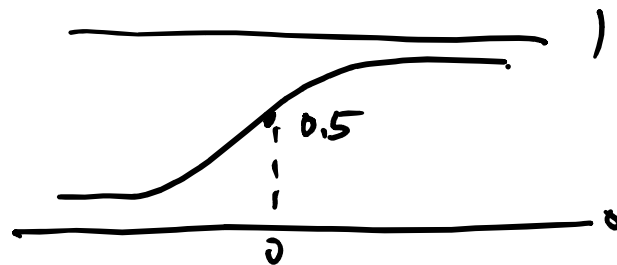
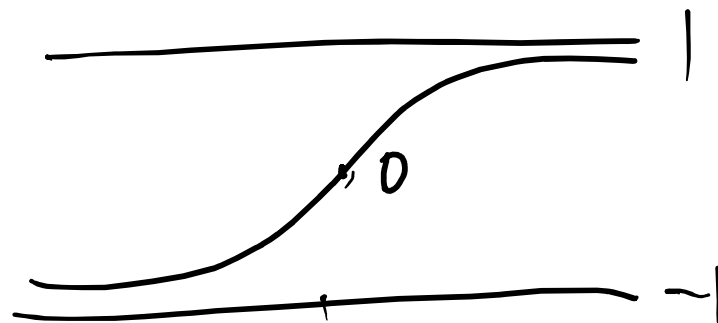


· RNN/LSTM.

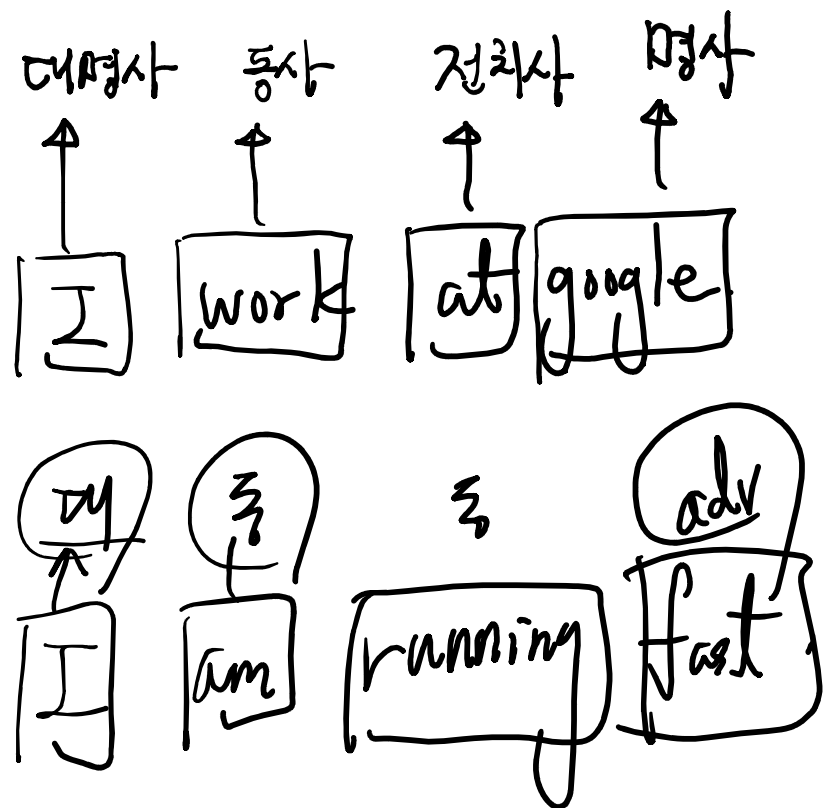
- σ (Sigmoid).



- tanh.



- Sequence

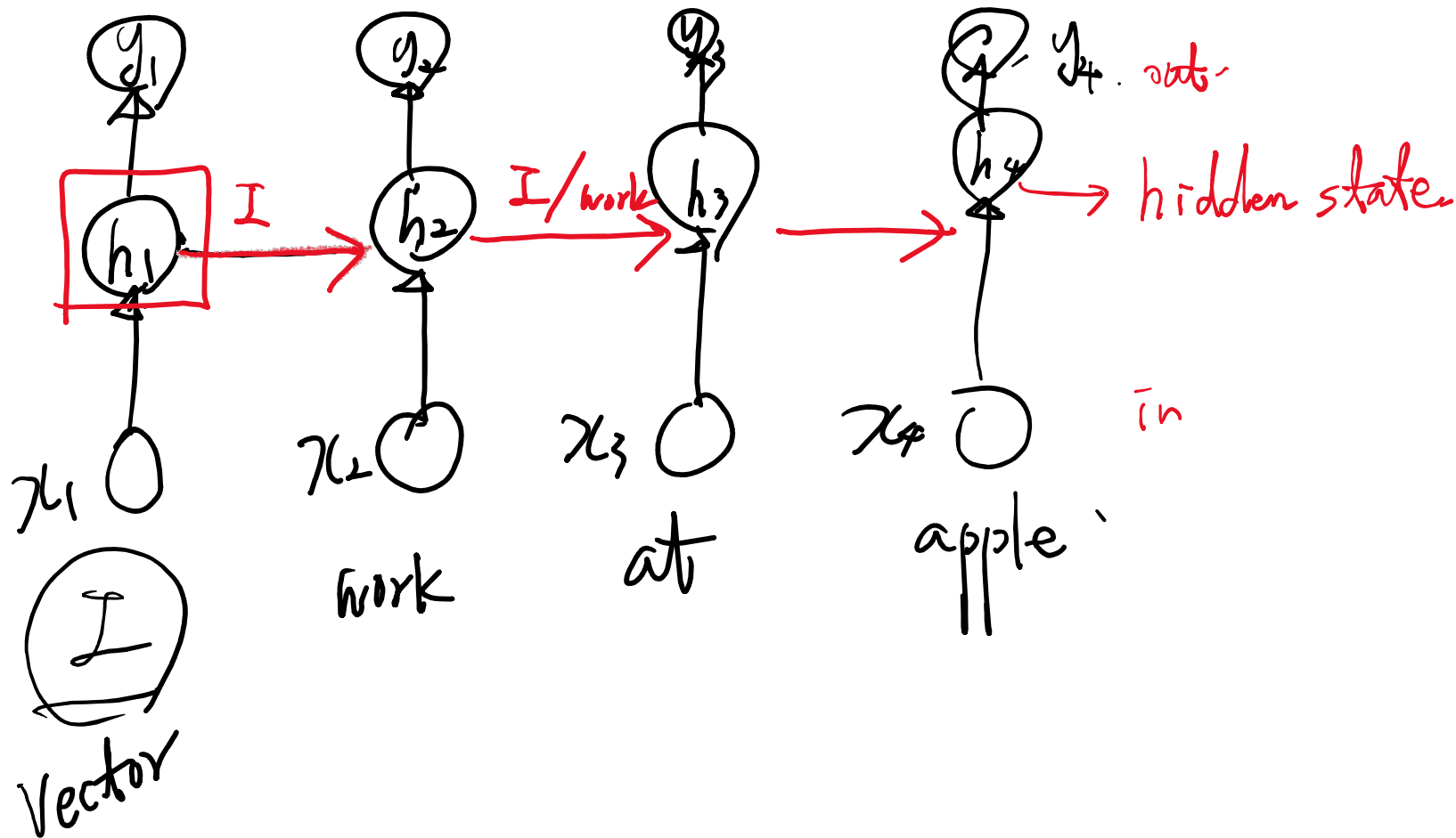


(part of speech: POS).

⇒ Supervised Learning :
POS tagging

I →	eat	apple.
I →	work	at apple.
t	$t + \alpha$	$t + \beta$

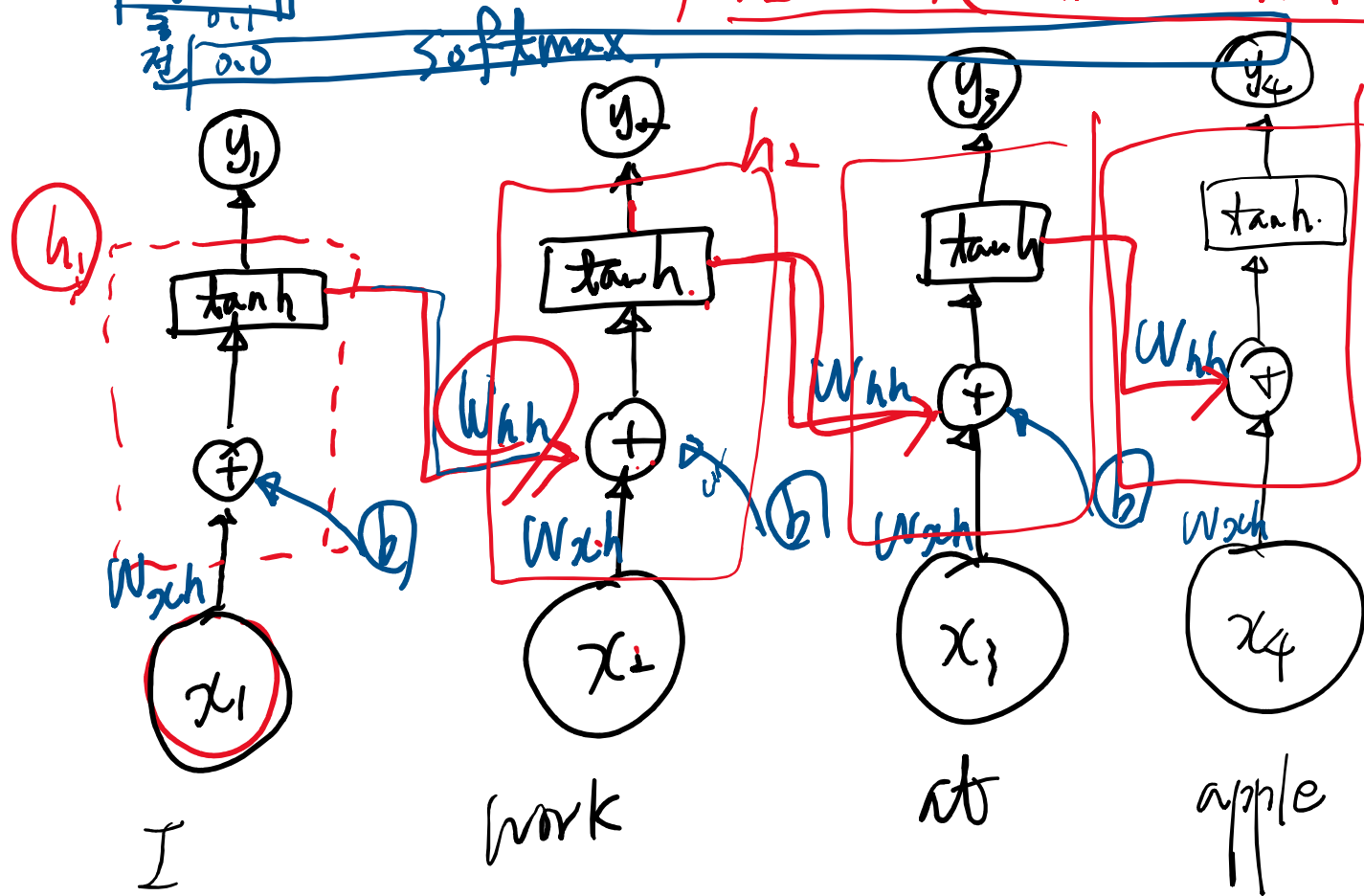
시간적, 연속성을 가지는 Data를
처리하는 DMV(DL) = RNN
↓



x_1	0.8
x_2	0.9
x_3	0.1
x_4	0.0

$$h_2 = \tanh(W_{xh} \cdot x_2 + W_{hh} h_1 + b)$$

$$W' = W - \eta \frac{\partial E}{\partial W} \quad y = Wx + b$$



bp.
bp through time
bp to t

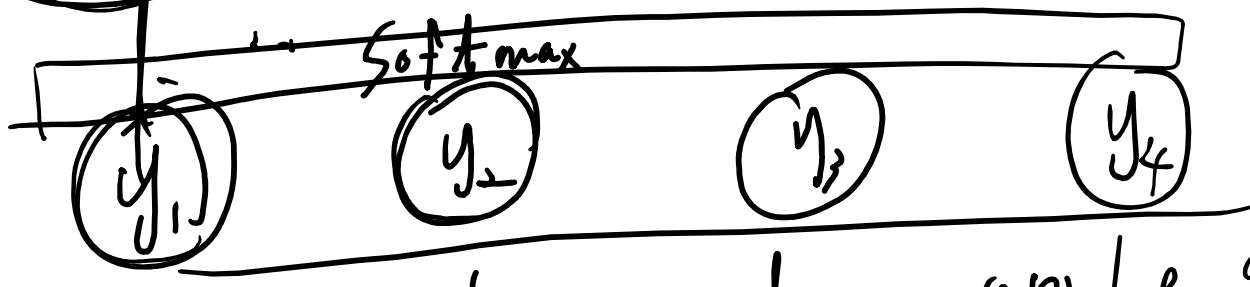
L_f
 prediction
 P_{old}
 P_1

040x4	0.8
03x4	0.1
5x4	0.1
21314	0.0

Label $\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$
 E_1

$$E_{error} = \sum_i E_i \quad (i=4)$$

avg min E
 θ

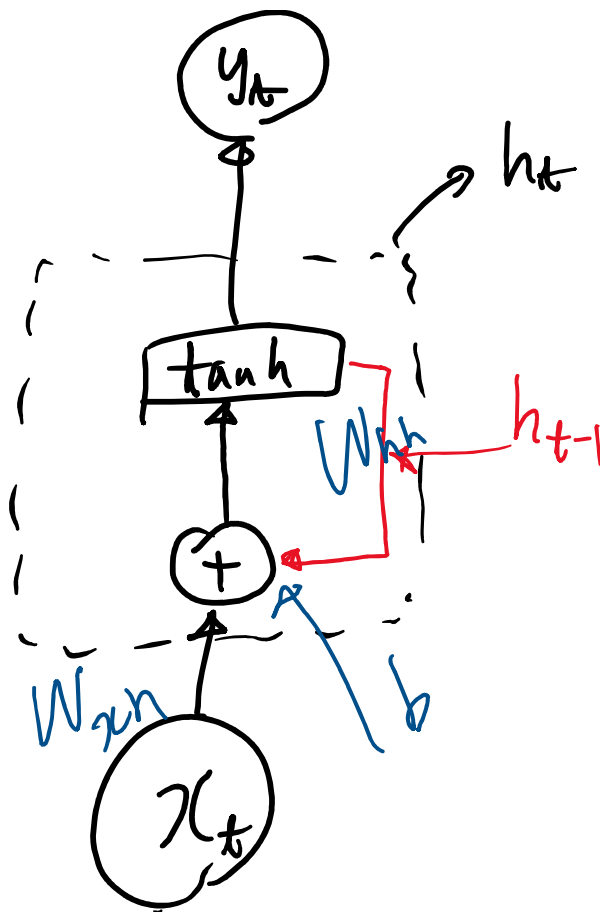


I work at apple

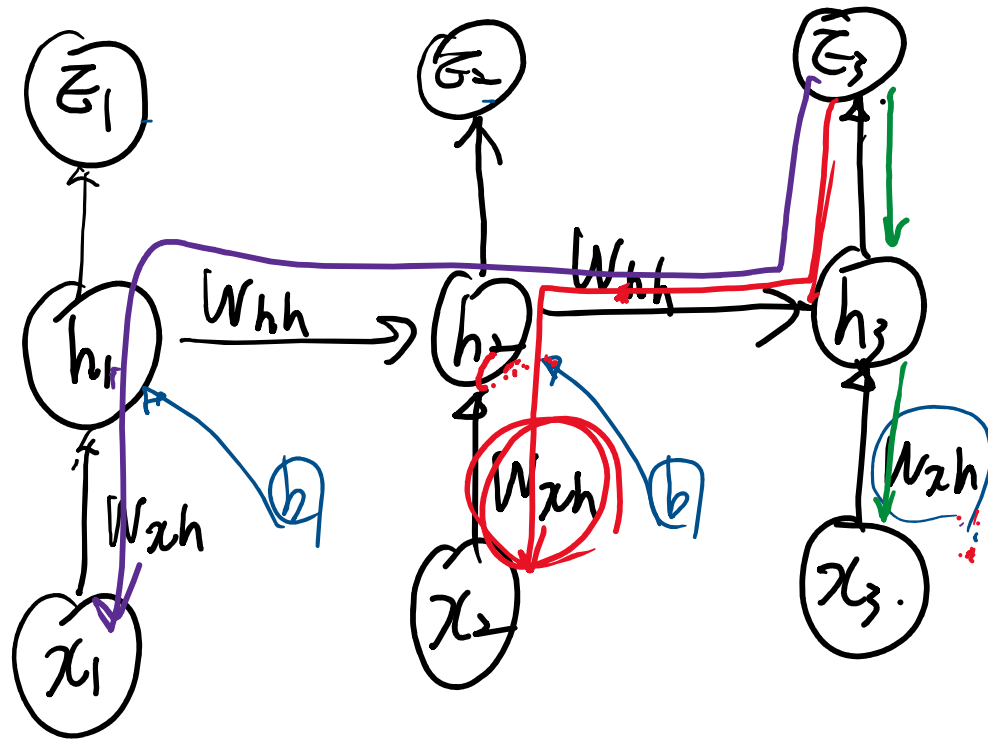
$$\theta = \begin{bmatrix} W_{zh} \\ W_{bh} \\ b \end{bmatrix}$$

$$W' = W - \eta \frac{\partial E}{\partial W}$$

Generalization



$$h_t = \tanh \left[W_{xh} \cdot x_t + W_{hh} \cdot h_{t-1} + b \right]$$

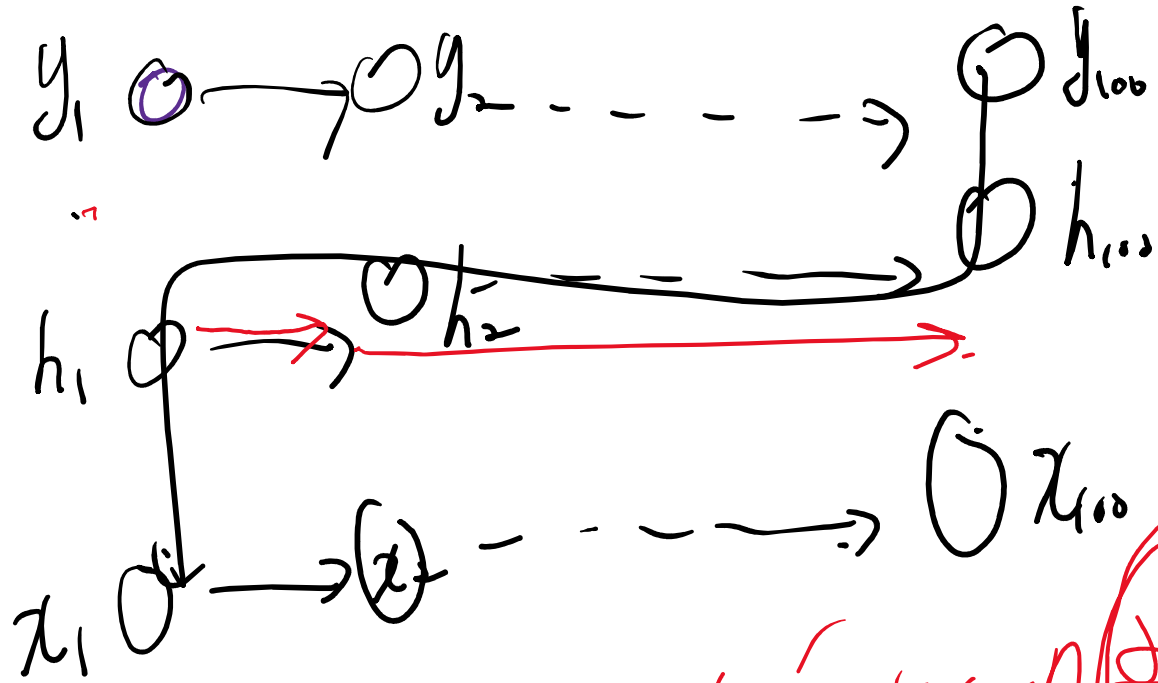


$$i) \left(\frac{\partial E}{\partial W} \right) = \frac{\partial E_1}{\partial W} + \frac{\partial E_2}{\partial W} + \left(\frac{\partial E_3}{\partial W} \right)$$

$$ii) \frac{\partial E_3}{\partial W} = \frac{\partial E_3}{\partial h_3} \cdot \frac{\partial h_3}{\partial W_{xh}}$$

$$+ \frac{\partial E_3}{\partial h_3} \cdot \frac{\partial h_3}{\partial h_2} \cdot \frac{\partial h_2}{\partial W_{xh}}$$

$$+ \frac{\partial E_3}{\partial h_3} \cdot \frac{\partial h_3}{\partial h_2} \cdot \frac{\partial h_2}{\partial h_1} \cdot \frac{\partial h_1}{\partial W_{xh}}$$



$$\frac{\partial E_{100}}{\partial w} \frac{\partial E_{100}}{\partial h_{100}} \frac{\partial h_{100}}{\partial h_{99}} \dots \frac{\partial h_2}{\partial h_1} \frac{\partial h_1}{\partial y_1}$$

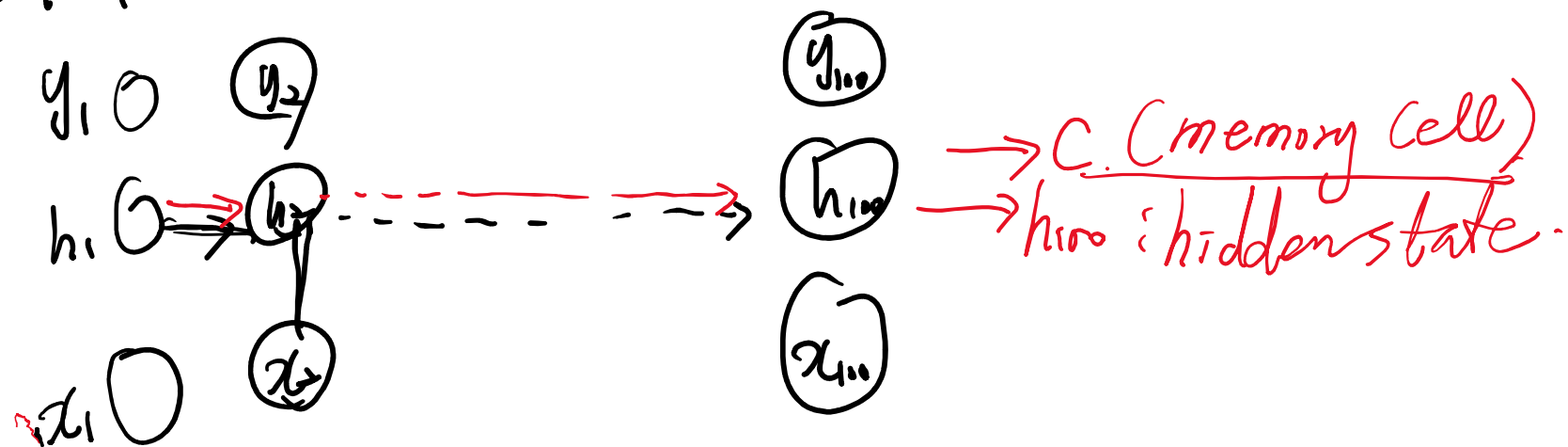
*if $\delta_i \Rightarrow \approx 0$
(gradient vanishing)*

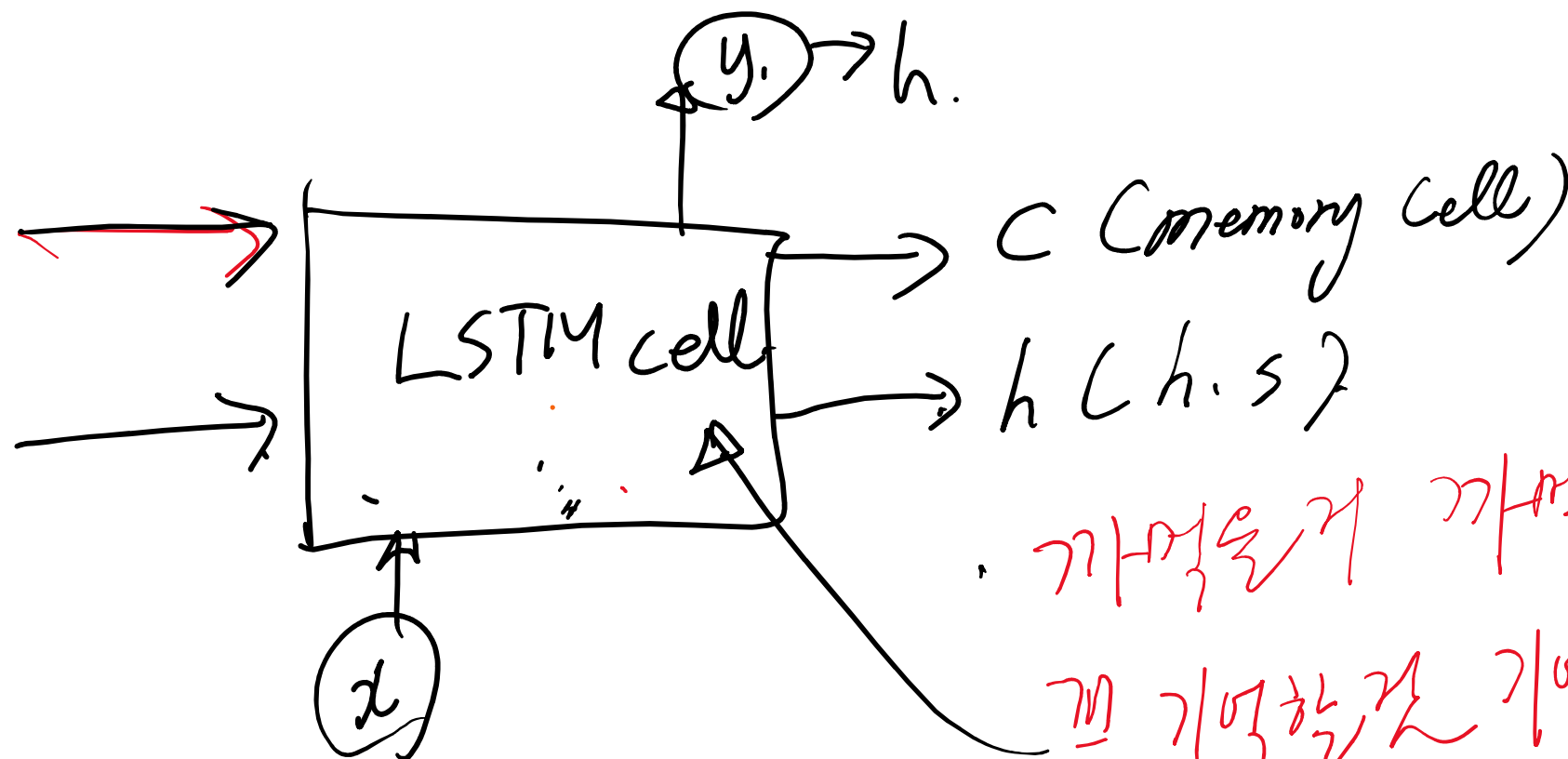
$$w' = w - \eta \frac{\partial E}{\partial w} \quad \text{if } \neq 0$$

else if \neq



LSTM





가장 먼저 기억하라
가장 나중에 기억하라

물찬 is my best friend. - :-,

물찬 is still my best friend.

He is still my best friend.

Jane is his g.f.

She knows I am best friend of him.

forget gate
input gate
output

