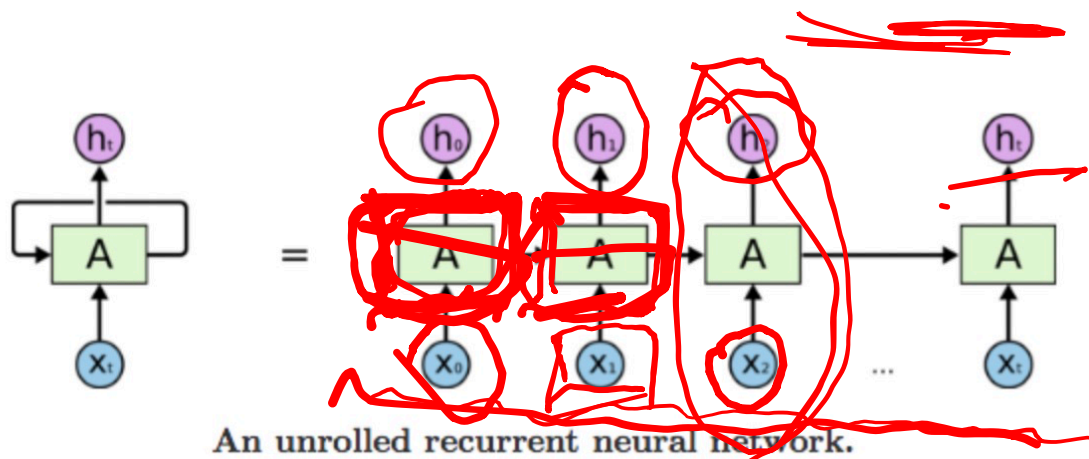
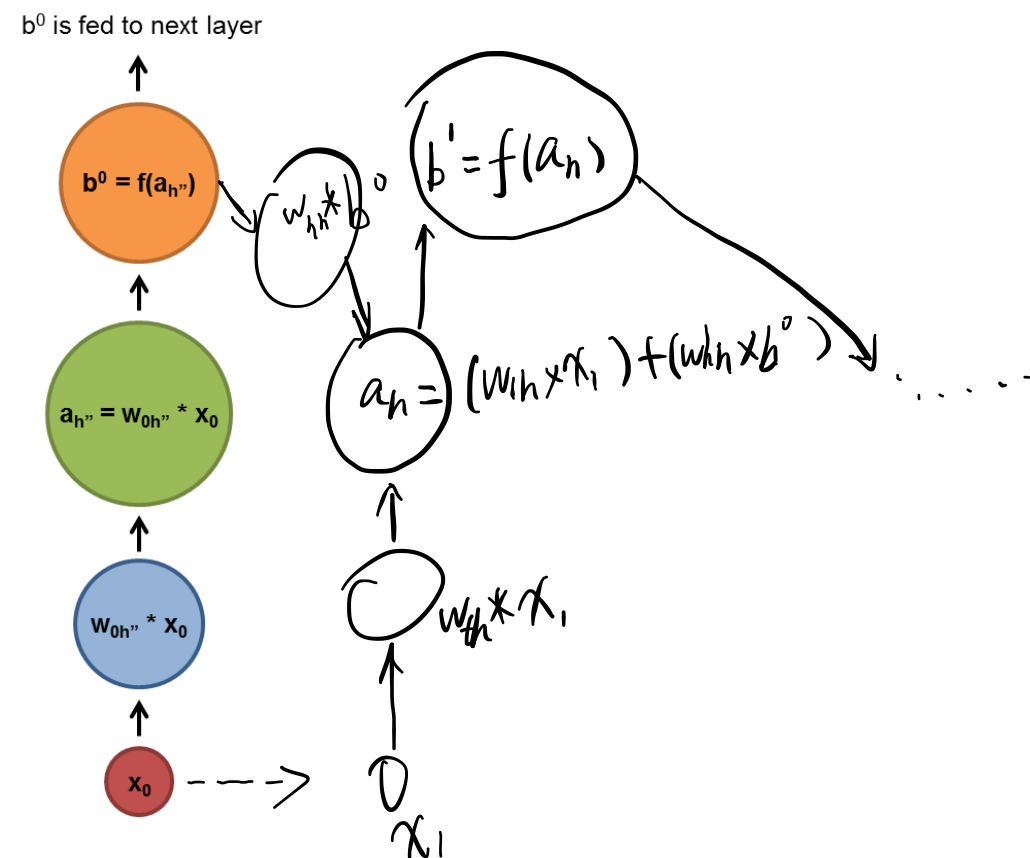


RNN

Recurrent Neural Network



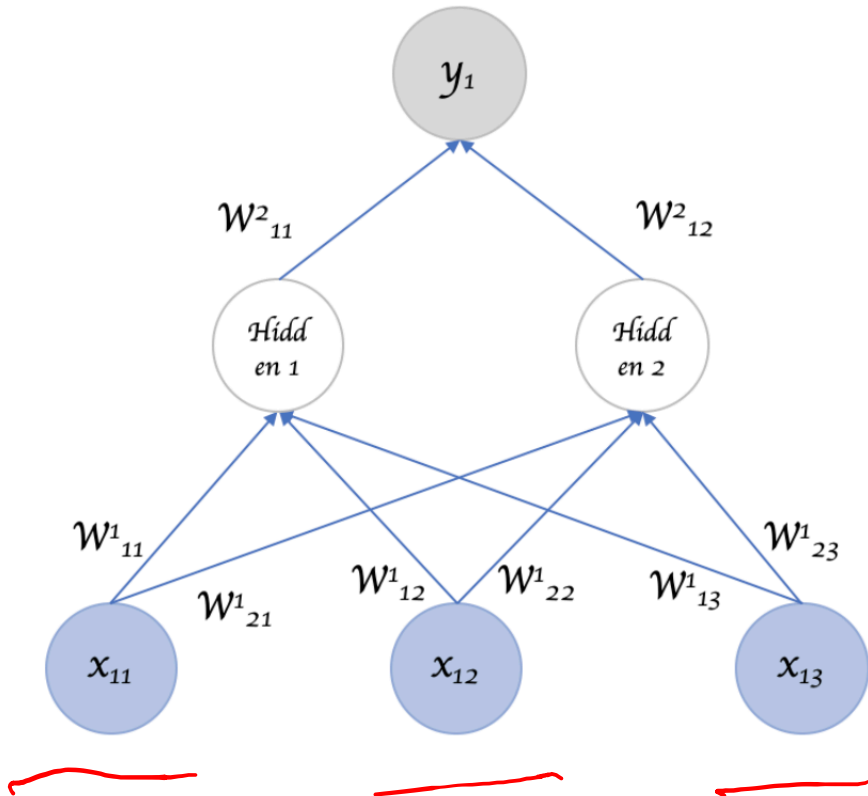
$$\underline{h_t = f(h_{t-1}, x_t)} \quad h_t = f(h_{t-1}, x_t)$$



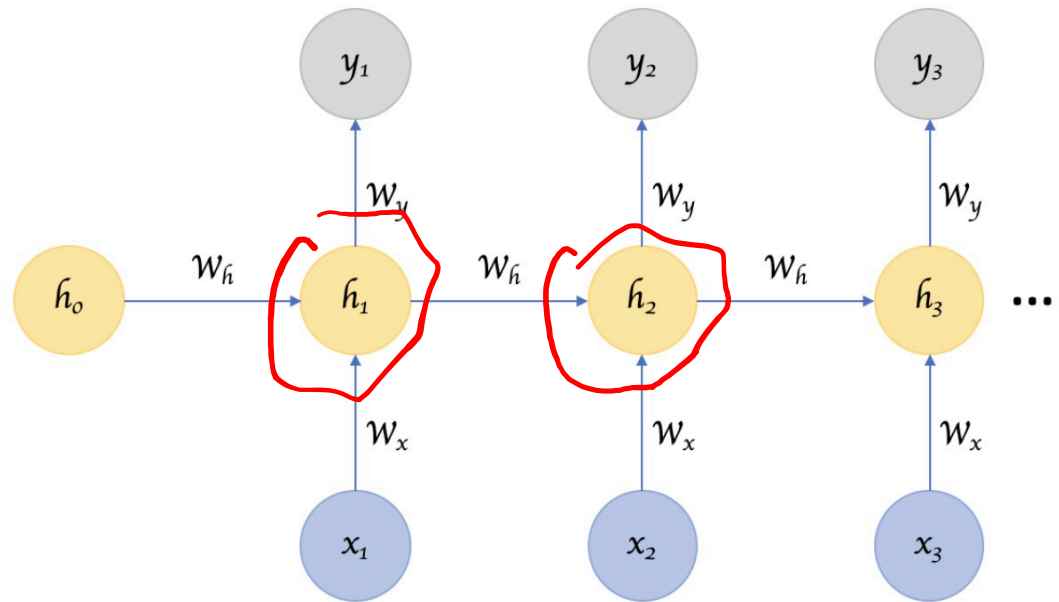
RNN

parameter sharing

CNN

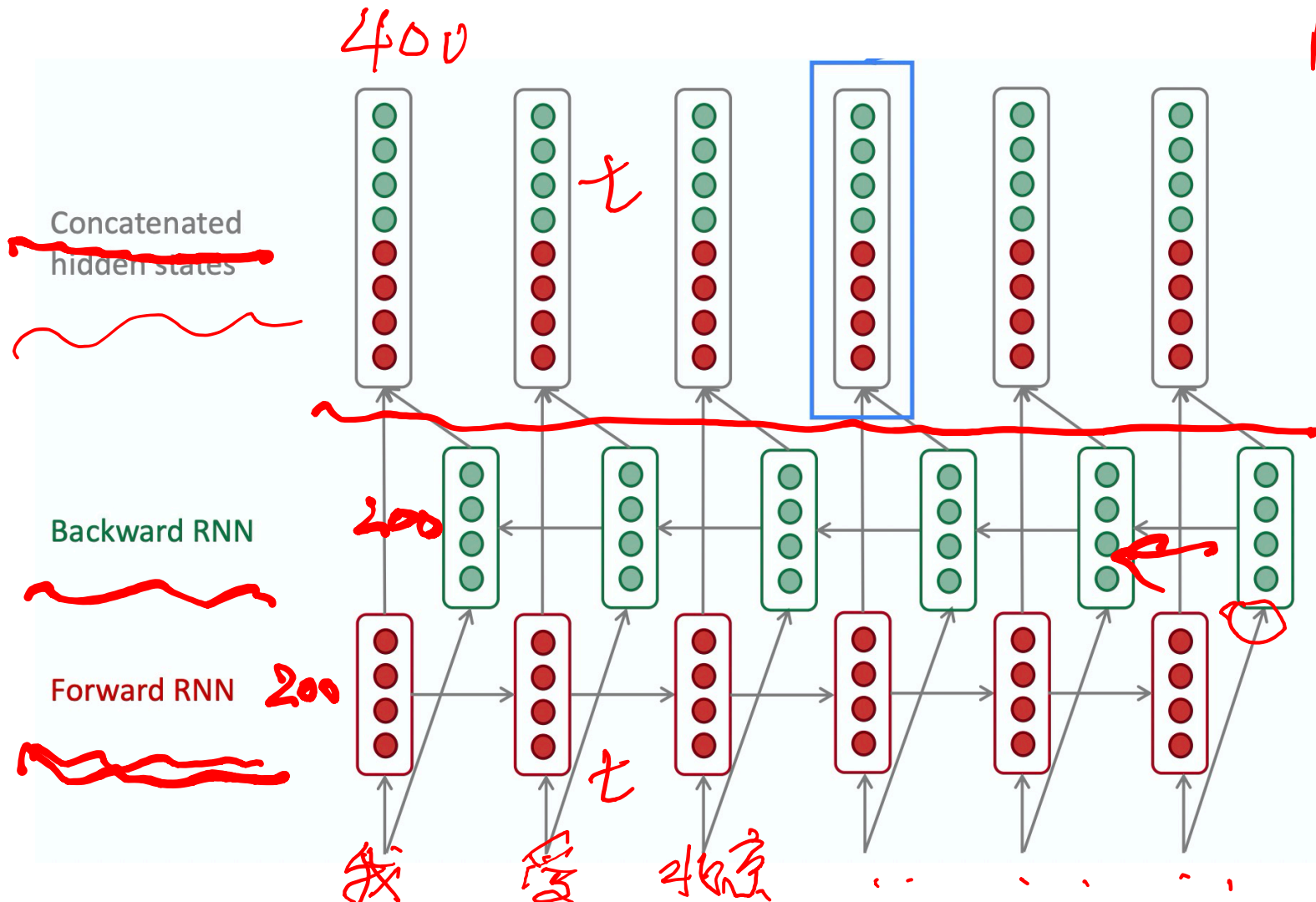


RNN



RNN

Bidirectional RNNs



$$F: \vec{h}^{(t)} = \text{RNN}_{\text{FW}}(\vec{h}^{(t-1)}, x^{(t)})$$

$$B: \overleftarrow{h}^{(t)} = \text{RNN}_{\text{BW}}(\overleftarrow{h}^{(t+1)}, x^{(t)})$$

$$h^{(t)} = [\vec{h}^{(t)}; \overleftarrow{h}^{(t)}]$$

RNN

Deep RNNs

$\left[\begin{bmatrix} 1, 2 \end{bmatrix}, \begin{bmatrix} 3, 4 \end{bmatrix} \right]$ $\xrightarrow{\text{axis} = 1}$ NMT
Network

