

RNN 基本原理:

Recall: 上禮拜的 NN (Simple)

$$\hat{h}_t = W_t X_t + b_t$$

$$\hat{y}_t = A(\hat{h}_t), \text{ where } A(\cdot) \text{ is the activation function.}$$

注意到, 這樣的 model 並沒有考慮到前一期  $h_{t-1}$ .

若真實的  $h_t$  和  $h_{t-1}$  有關 (e.g.  $h_t = f(h_{t-1})$ )

則模型預測會出現問題.

eg: 金融資料 (當天股價和前幾天股價有關)  
語意判斷 (字序順序導致不同意義)

因此, 我們優化了 NN, 使其能夠考慮前一期的結果.

RNN formula:

$$h_t = W_{it} \cdot h_{t-1} + W_{xt} \cdot X_t + b_t$$

$$y_t = A(h_t)$$

, where  $A(\cdot)$  is activation function.

Note:

事實上, RNN 不只考慮前一期的結果, 它考慮了  $t$  期之前的所有期數.

Since

$$h_1 = h_1$$

$$h_2 = W_{12}h_1 + W_{21}x_1 + b_2$$

$$h_3 = W_{13}h_2 + W_{23}x_2 + b_3$$

$$= W_{13}W_{12}h_1 + W_{13}W_{21}x_1 + W_{23}x_2 + \underbrace{b_3 + W_{13}b_2}_{z=b}$$

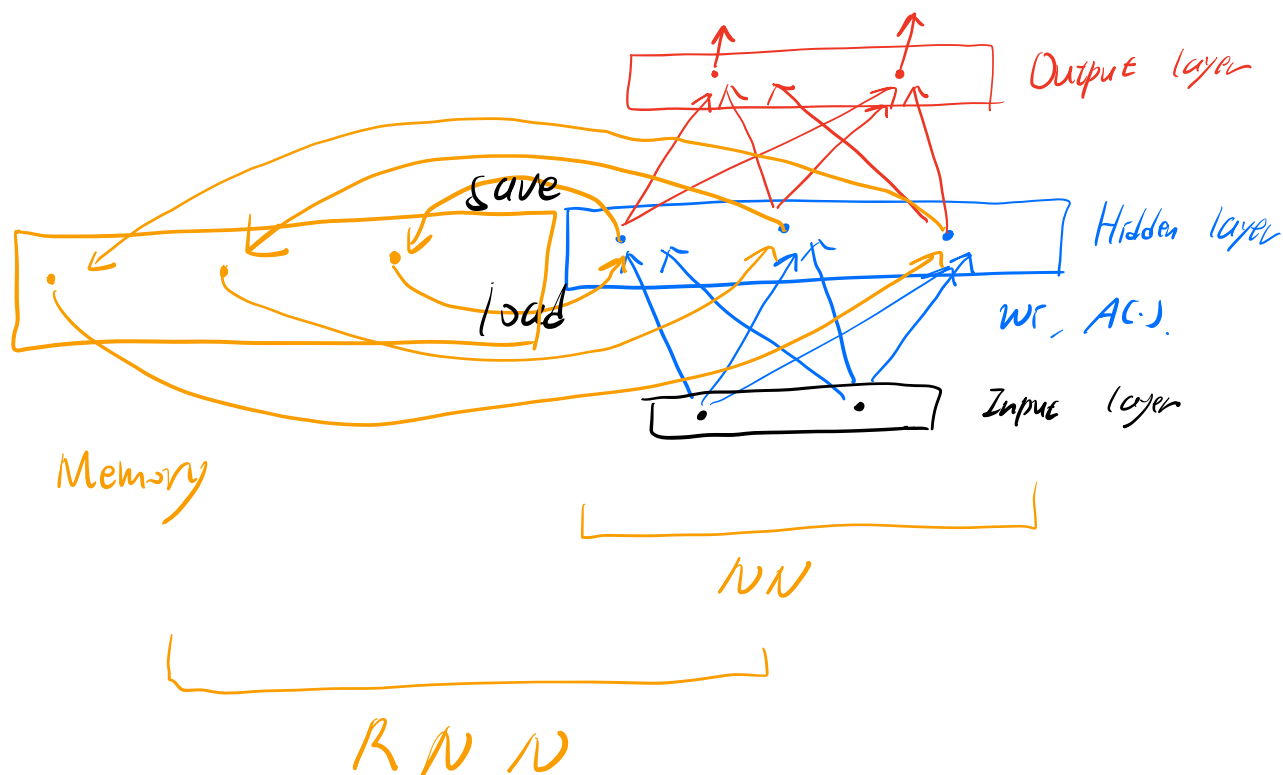
⋮

$$h_t = W_{1t}h_{t-1} + W_{2t}x_t + b_t$$

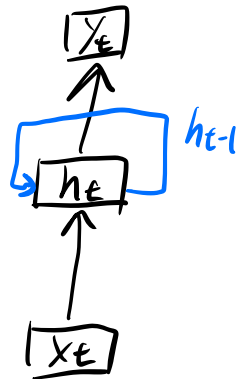
$$= W_{1t} \left( \sum_{k=2}^{K-1} W_{1k}h_1 + \sum_{k=2}^{K-1} W_{2k}x_k \right) + W_{2t}x_t + b_t$$

和前面全部有關!

畫成圖的話:



Simplified :



Refinement :

LSTM & GRU.  
(~~統稱~~ RNN)

Example :

IMBD

philips curve

SP 500