



Deep learning with keras

CH10

CONTENT

- 01** 各種時間序列任務
- 02** 溫度預測任務
- 03** 認識循環神經網路(RNN)
- 04** 循環神經網路的進階應用



10-1

各種時間序列任務

時間序列任務的種類

- 分類

ex : 依網路活動紀錄判斷人或機器人

- 事件偵測

ex : 熱詞偵測 (hotword detection)

- 異常偵測 (非監督學習)

ex : 網路上的異常活動、生產線上的異常狀況



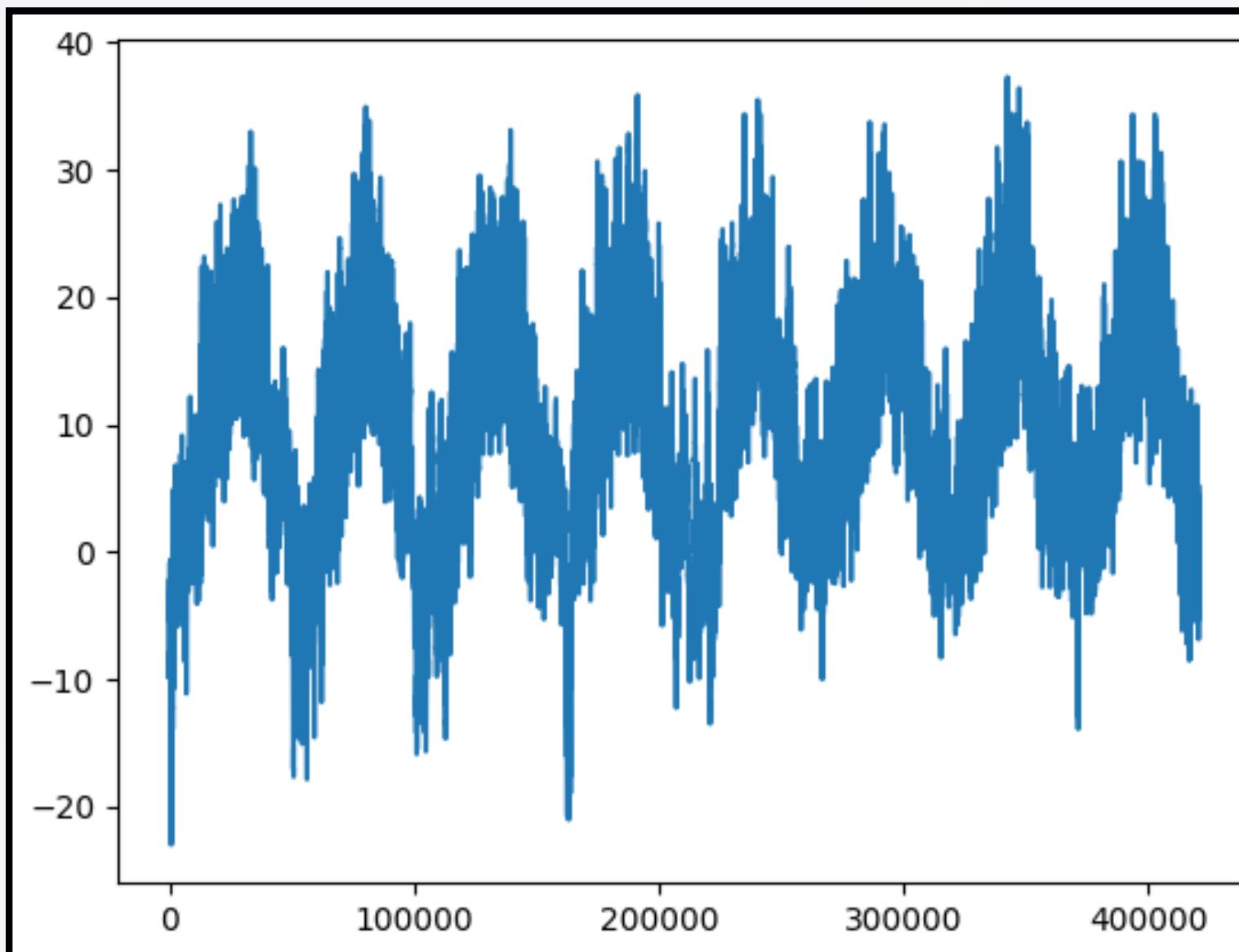
10 - 2

溫度預測任務

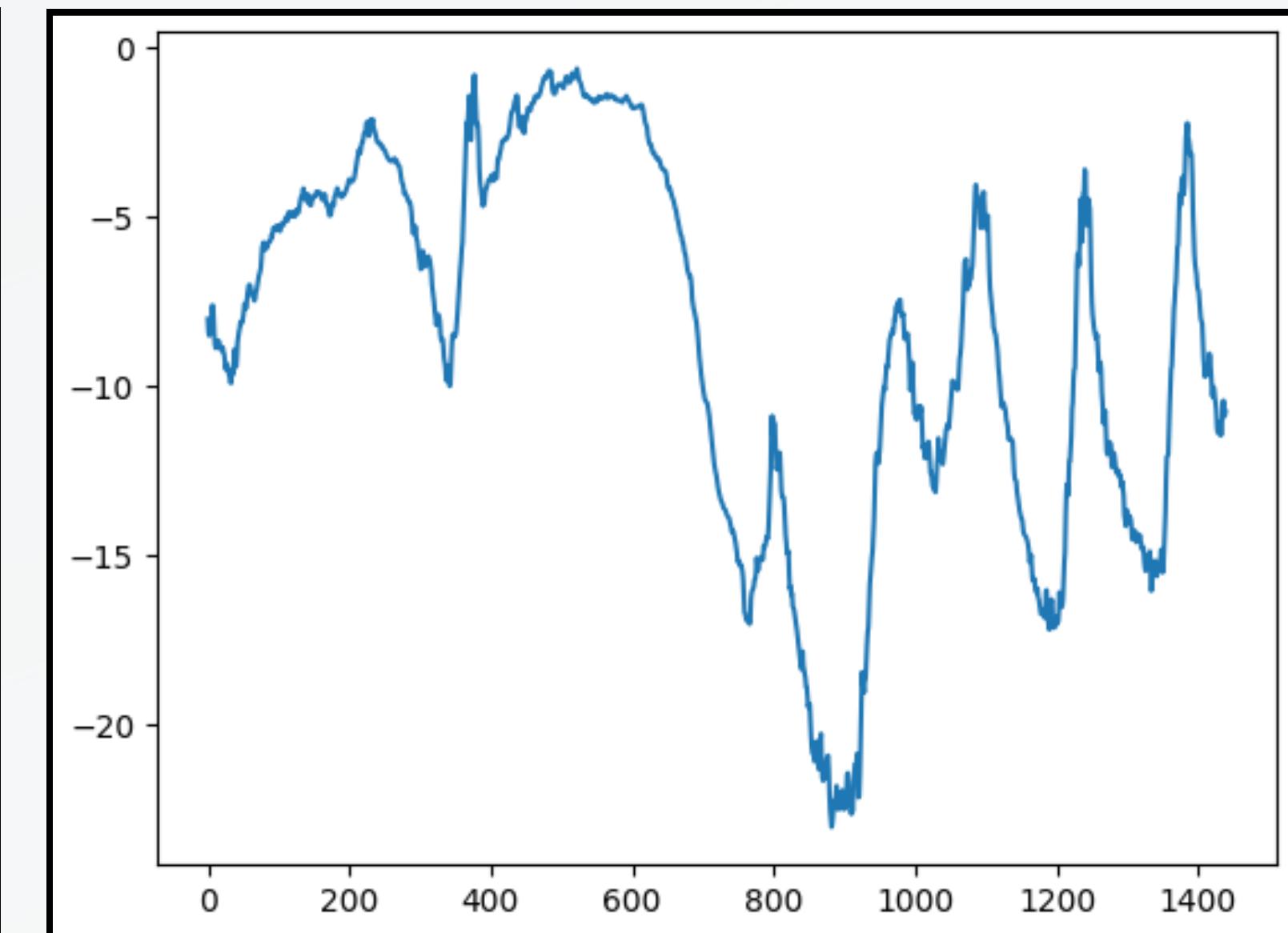
時間序列資料

- DNN 和 CNN 都不具備處理能力
- RNN 適用於該問題
- 以德國耶拿氣象站紀錄的天氣時間序列資料集為例，
每10分鐘為區間測量各項天氣指標，資料集記錄範圍
為 2009~2016

時間序列資料



資料集溫度變化



前10天的資料變化

look code

若給定前5天的資料(每小時採樣一次), 我們能否預測24小時的溫度？

準備資料

- 資料正規化 (平均值, 標準差)
- 使用 `times_dataset_from_array` 生成序列
- 計算基準線的平均絕對誤差(MEA)

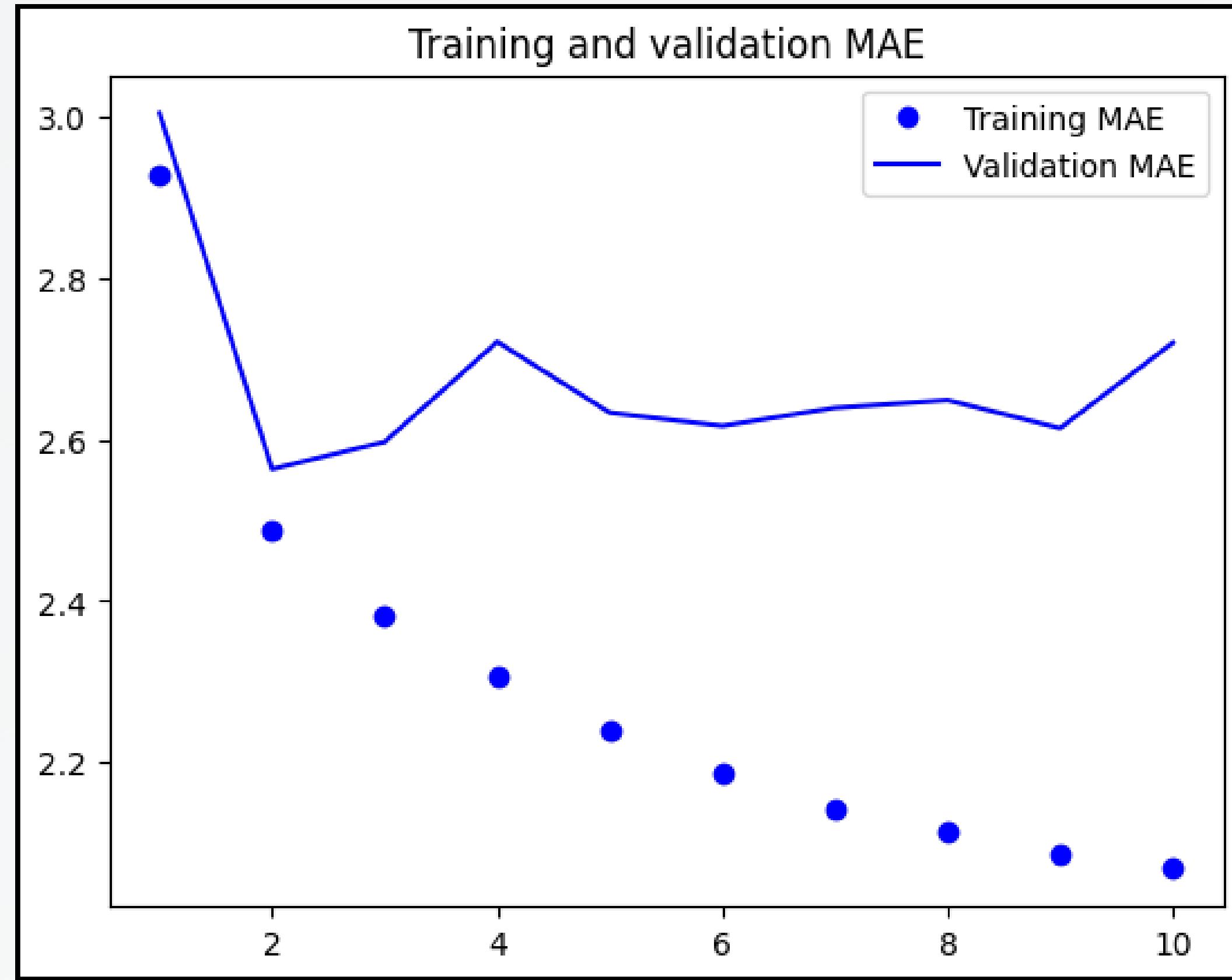
look code

密集連接層

- Final Dense layer 沒有激活函數 (回歸問題常用)
- 使用均方誤差(MSE)作為損失函數 (迎合梯度下降)
- MAE(2.62) 貼近基準線(2.44)
- 想找到假設模型, 等同於大海撈針

look code

密集連接層



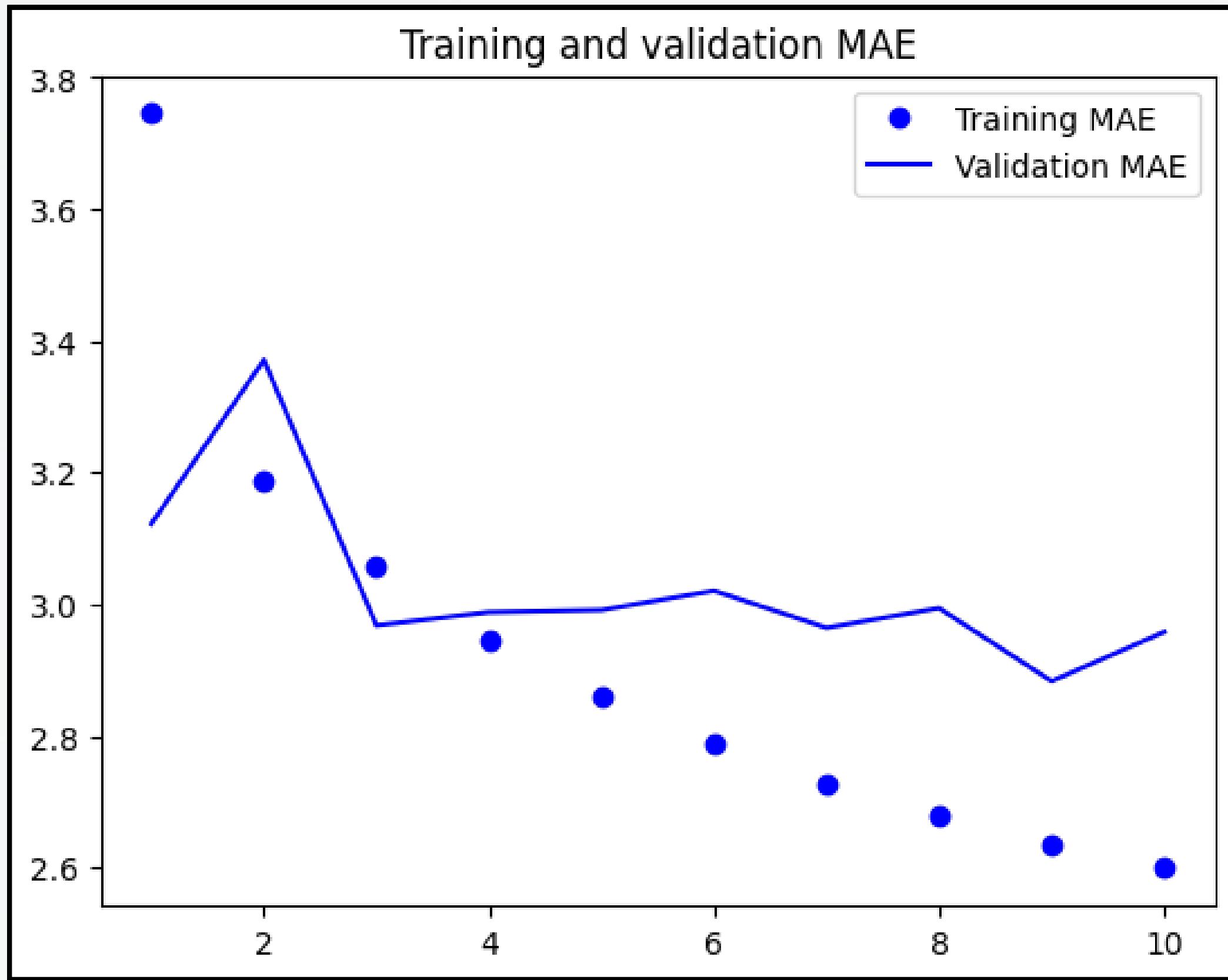
look code

1D 卷積模型

- 適用於具平移不變性的序列資料
- 資料具有以一天為一個循環的特點
- MAE(2.9) 與基準線(2.44)相差甚遠
- 早晚的資料特性不同
- 資料順序決定參考價值

look code

1D 卷積模型



look code

密集層 / 卷基層的問題

- 第一時間將時間序列扁平化, 喪失時間序
列資訊
- 池化層破壞時間序列資訊

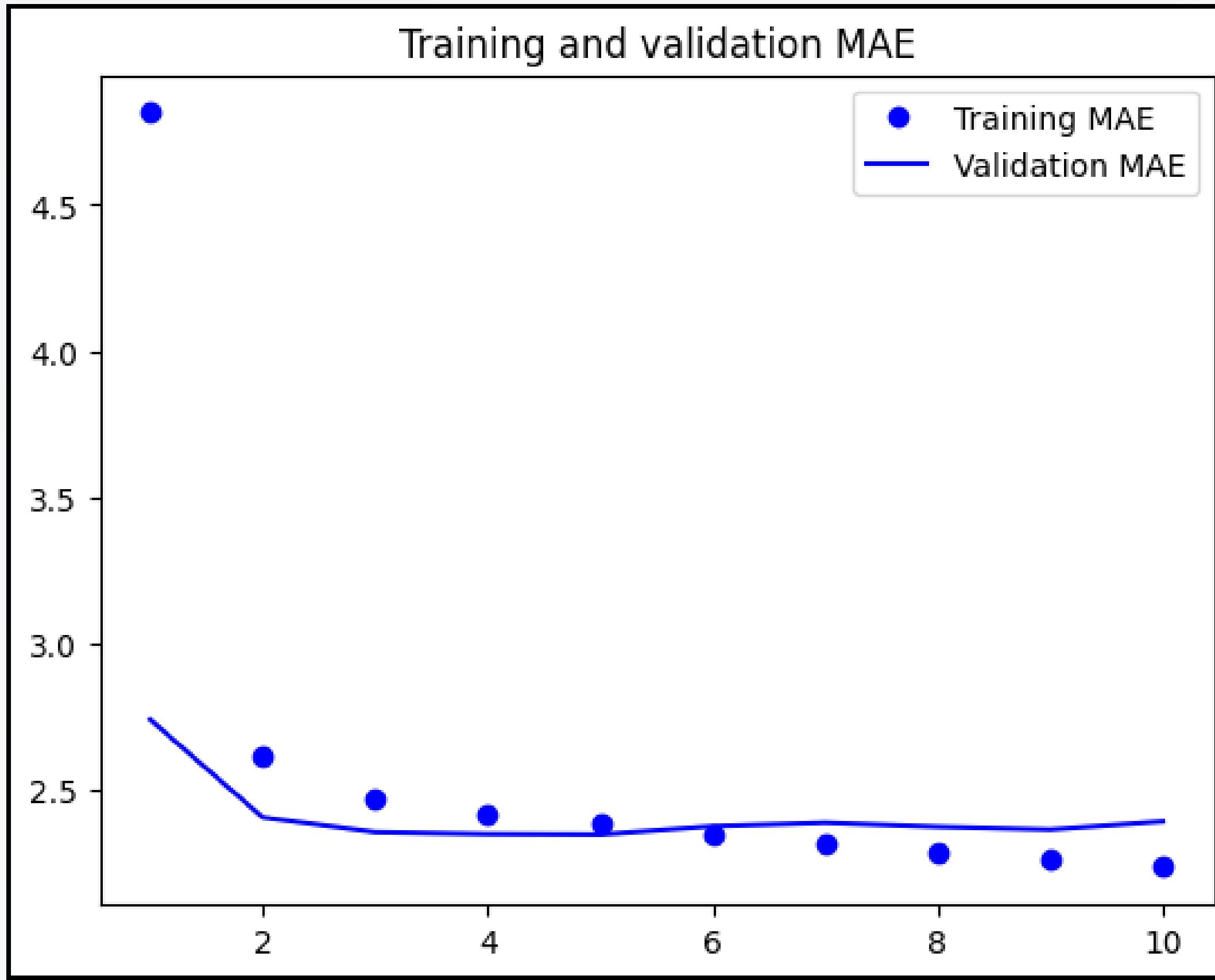
look code

循環神經網路

- 長短期記憶(LSTM) 最受歡迎
- MAE(2.35) 的成效最好
- MAE 超越基準線(2.44)

look code

1D 卷積模型



look code



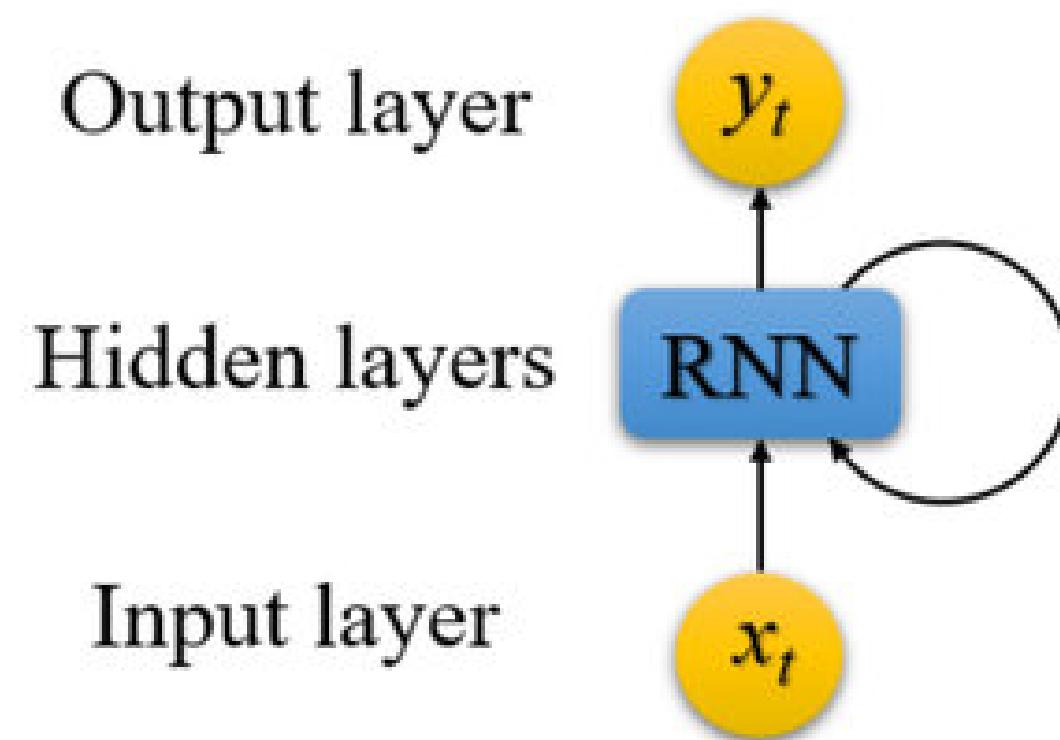
10 - 3

認識循環神經網路(RNN)

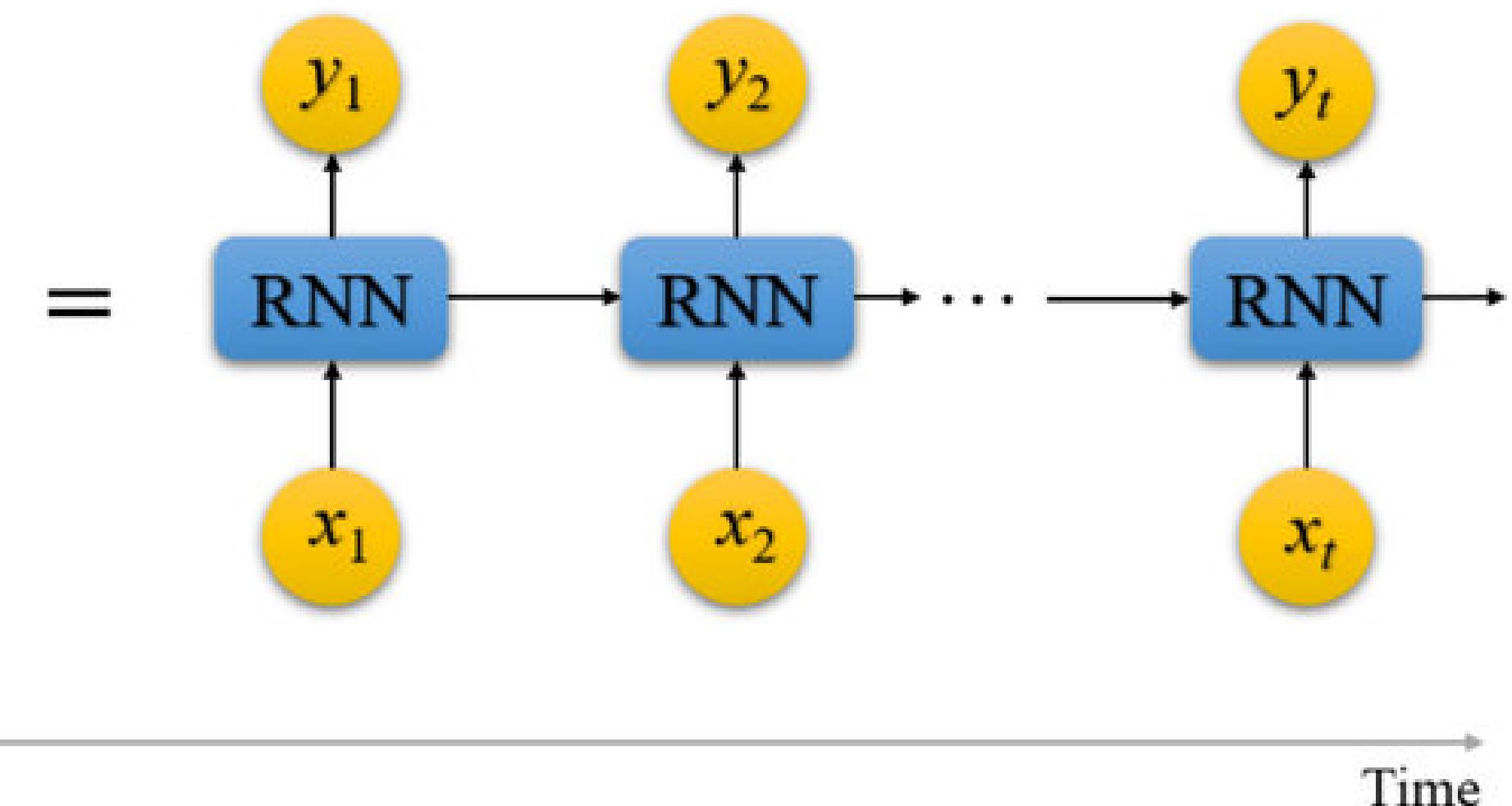
循環神經網路(RNN)

- 保留對先前內容的記憶
- 將一串時序當作一個資料點
- 對序列元素進行迭代

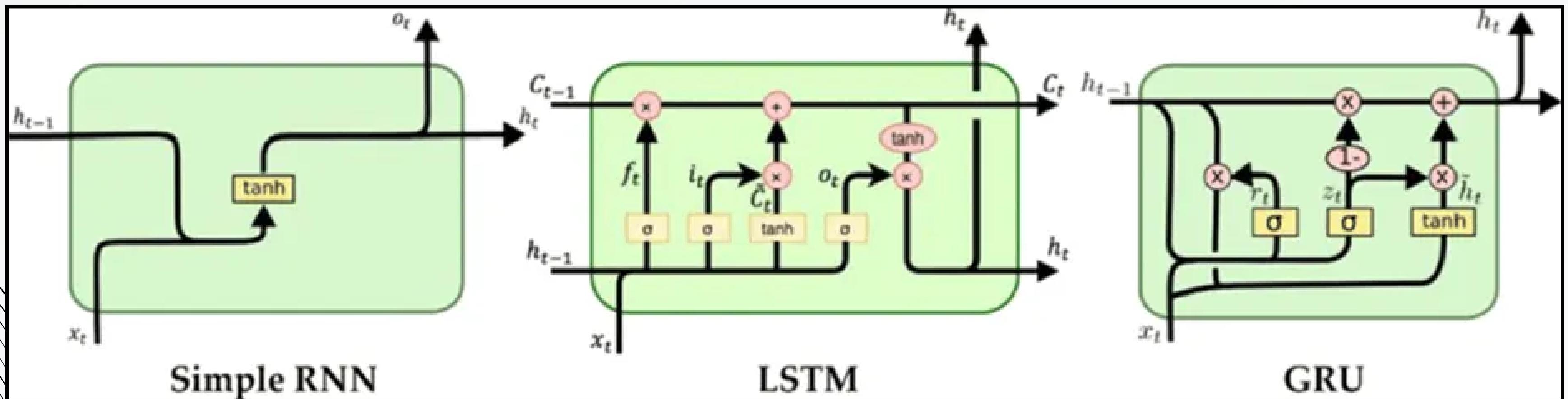
Rolled RNN



Unrolled RNN

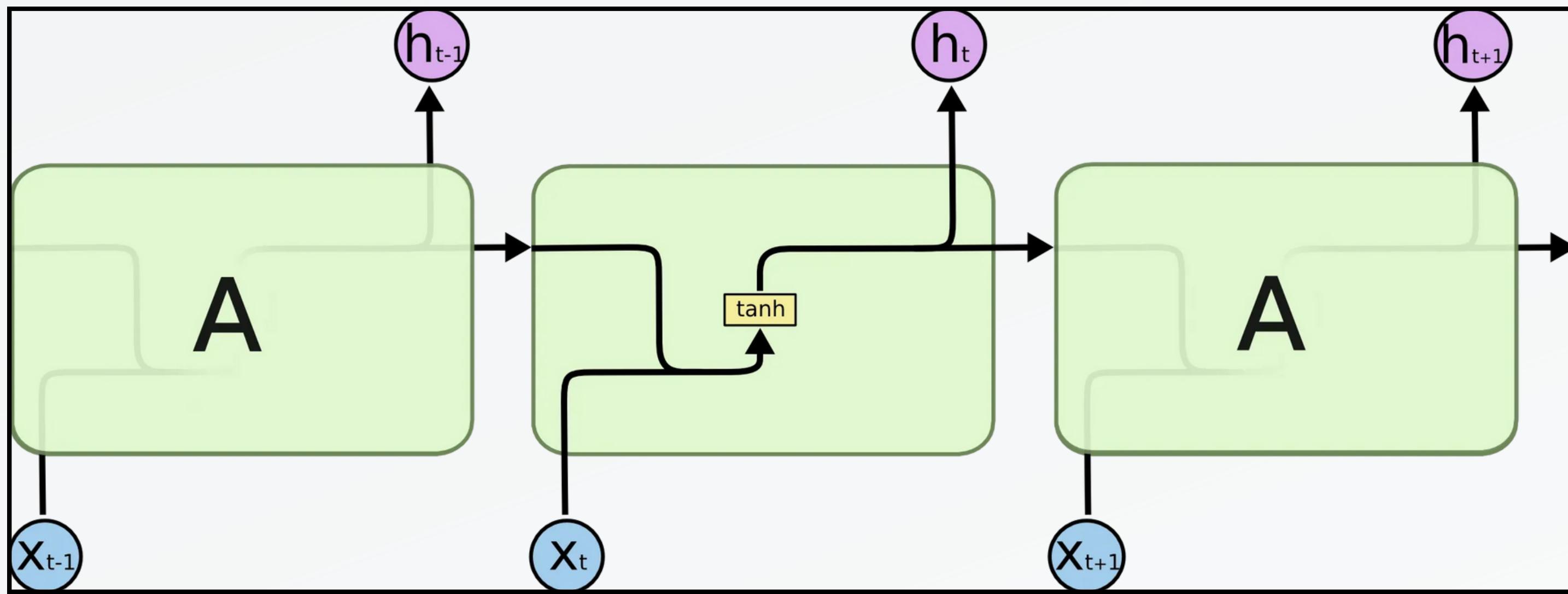


Keras 中的循環層



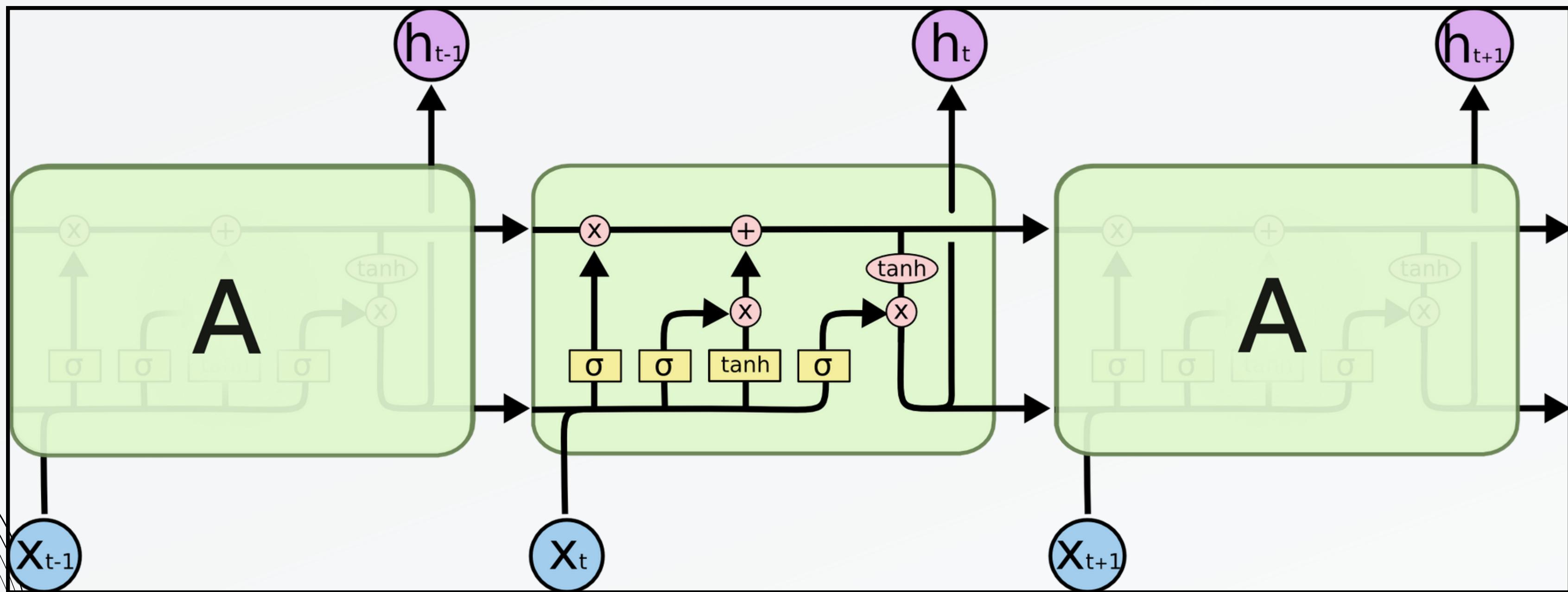
SimpleRNN

- 能夠以批次處理
- 可以處理任意長度的序列
- 存在梯度消失和梯度爆炸的問題
- 對於長期依賴性的捕捉能力有限



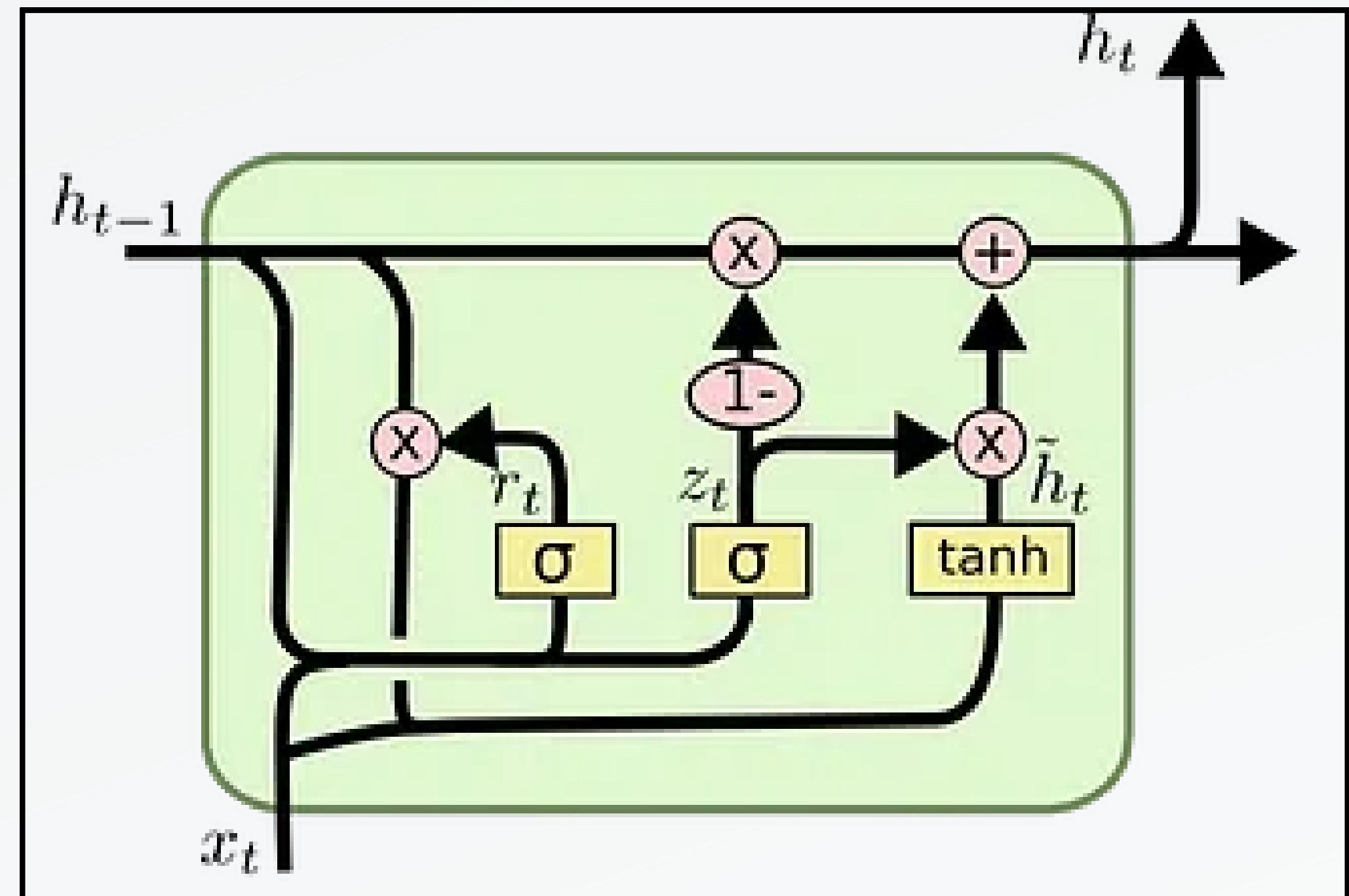
LSTM

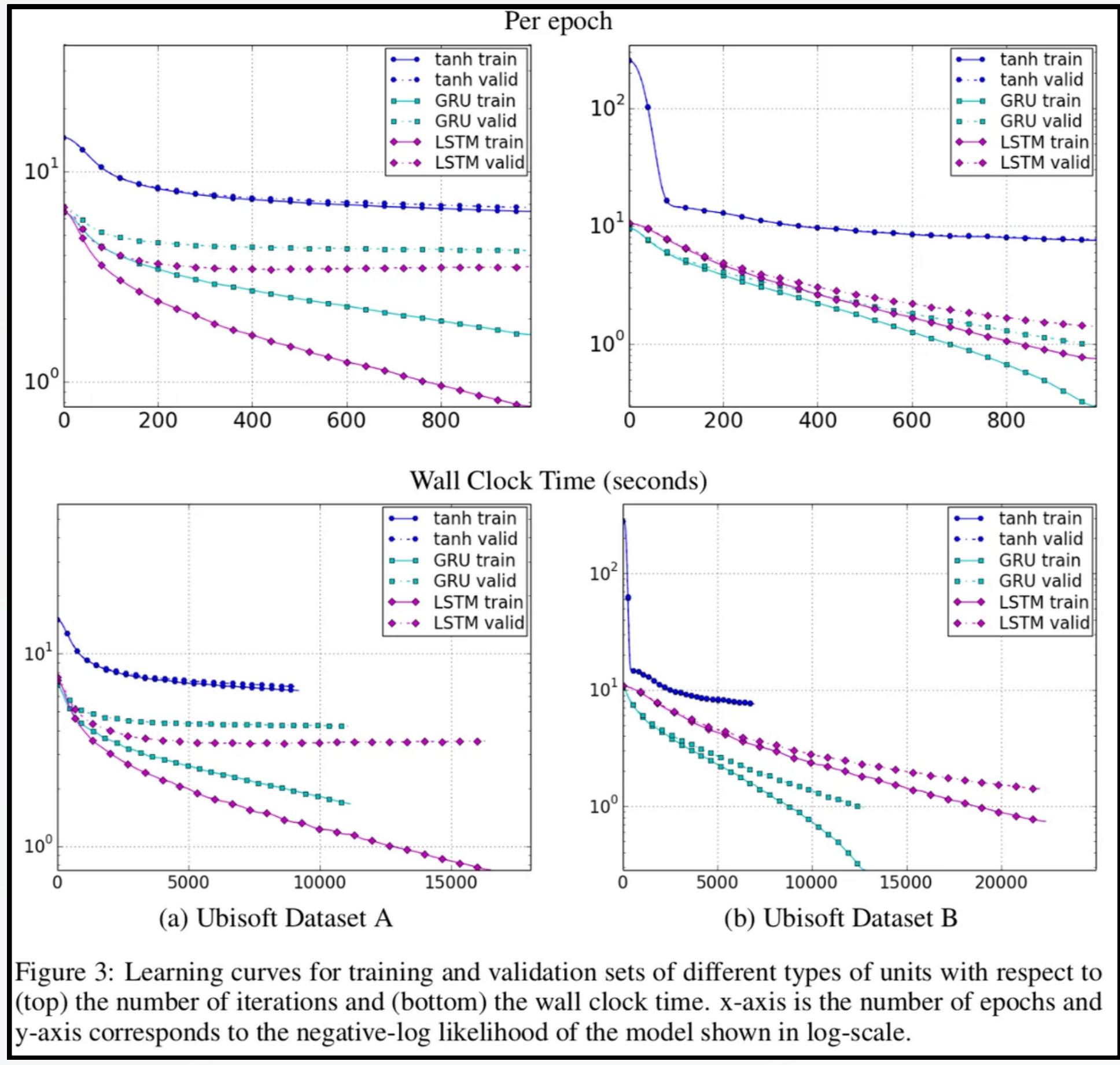
- 引入記憶單元
- 包含輸入門, 輸出門, 遺忘門
- 對於處理複雜的序列資料至關重要



GRU

- LSTM 的改良
- 執行速度快, 減少記憶體使用
- 以更新門取代遺忘和輸入門
- 記憶狀態和隱藏狀態合併





10-4

循環神經網路的進階應用

RNN的進階功能

- 循環丟棄法
- 堆疊多個 RNN
- 雙向 RNN

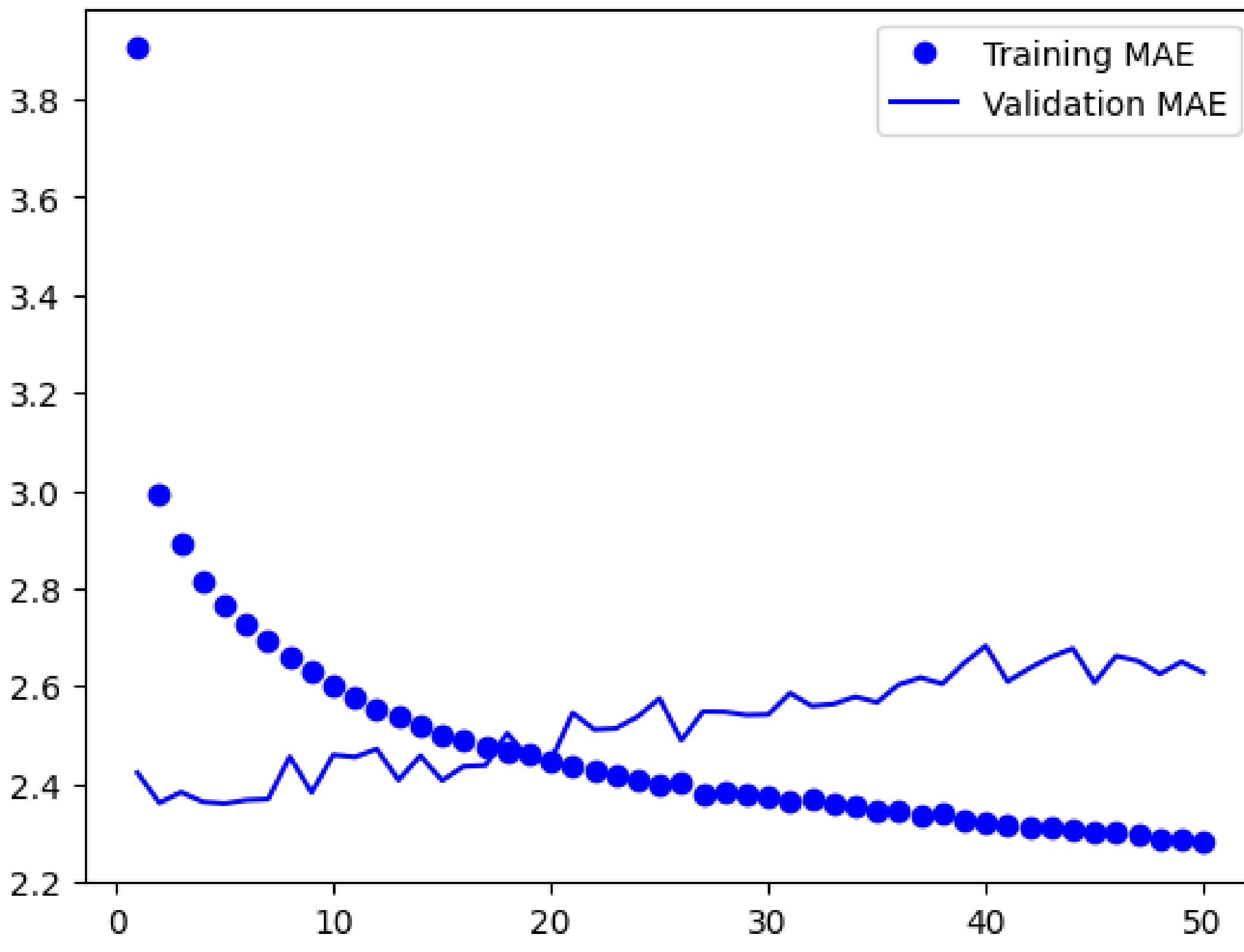
使用循環丟棄法對抗 overfit

- 丟棄法：隨機將部分輸入單元變為零
- 隨機變換(隨時間改變)的丟棄遮罩會阻礙學習
- 每個時步上，採用相同的丟棄遮罩

LSTM 加入循環丟棄法

- 使用雙倍單元數(32)
- 使用5倍周期數訓練(使用丟棄法要增加訓練)
- MAE(2.27)

Training and validation MAE

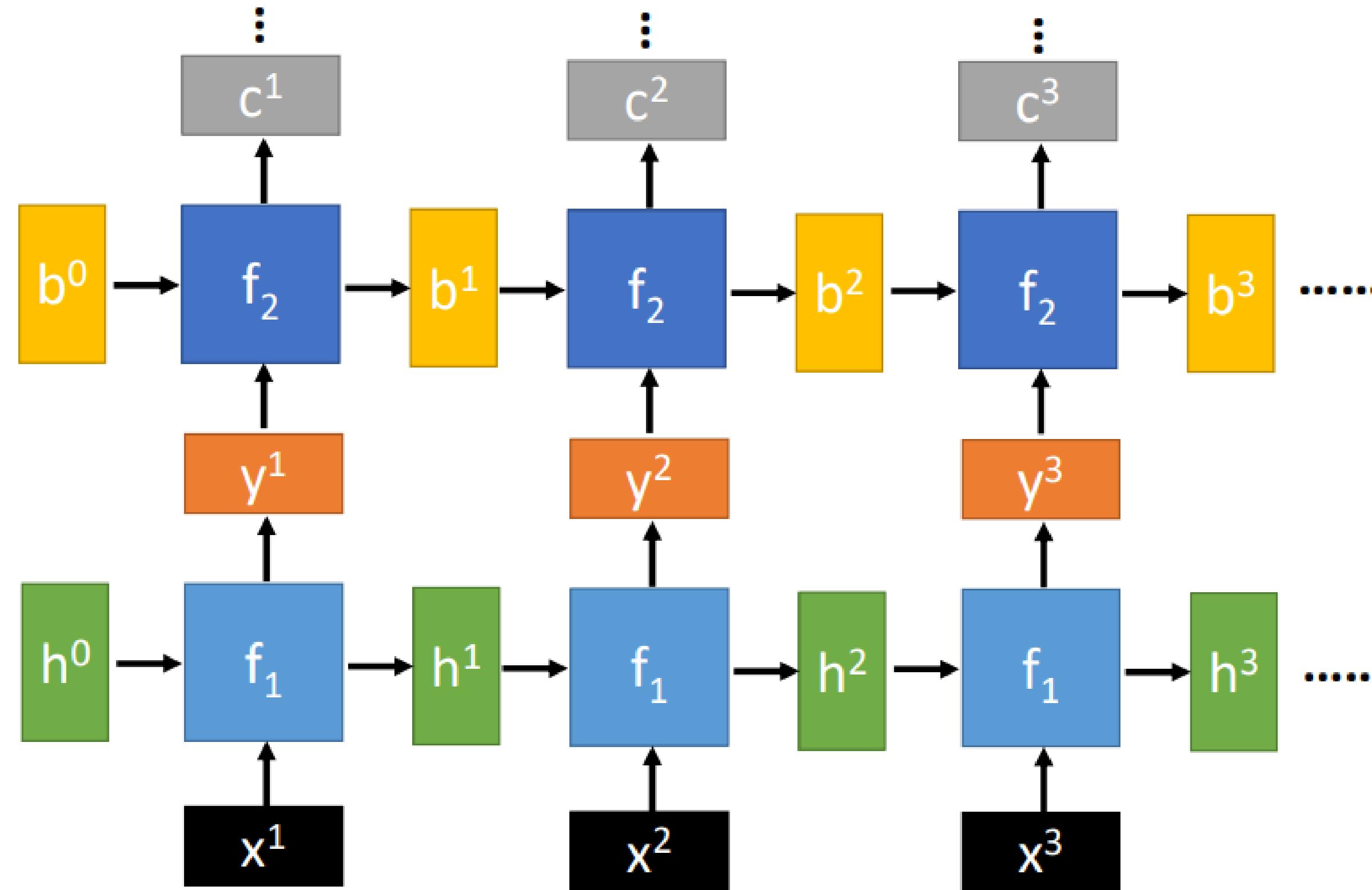


堆疊 RNN

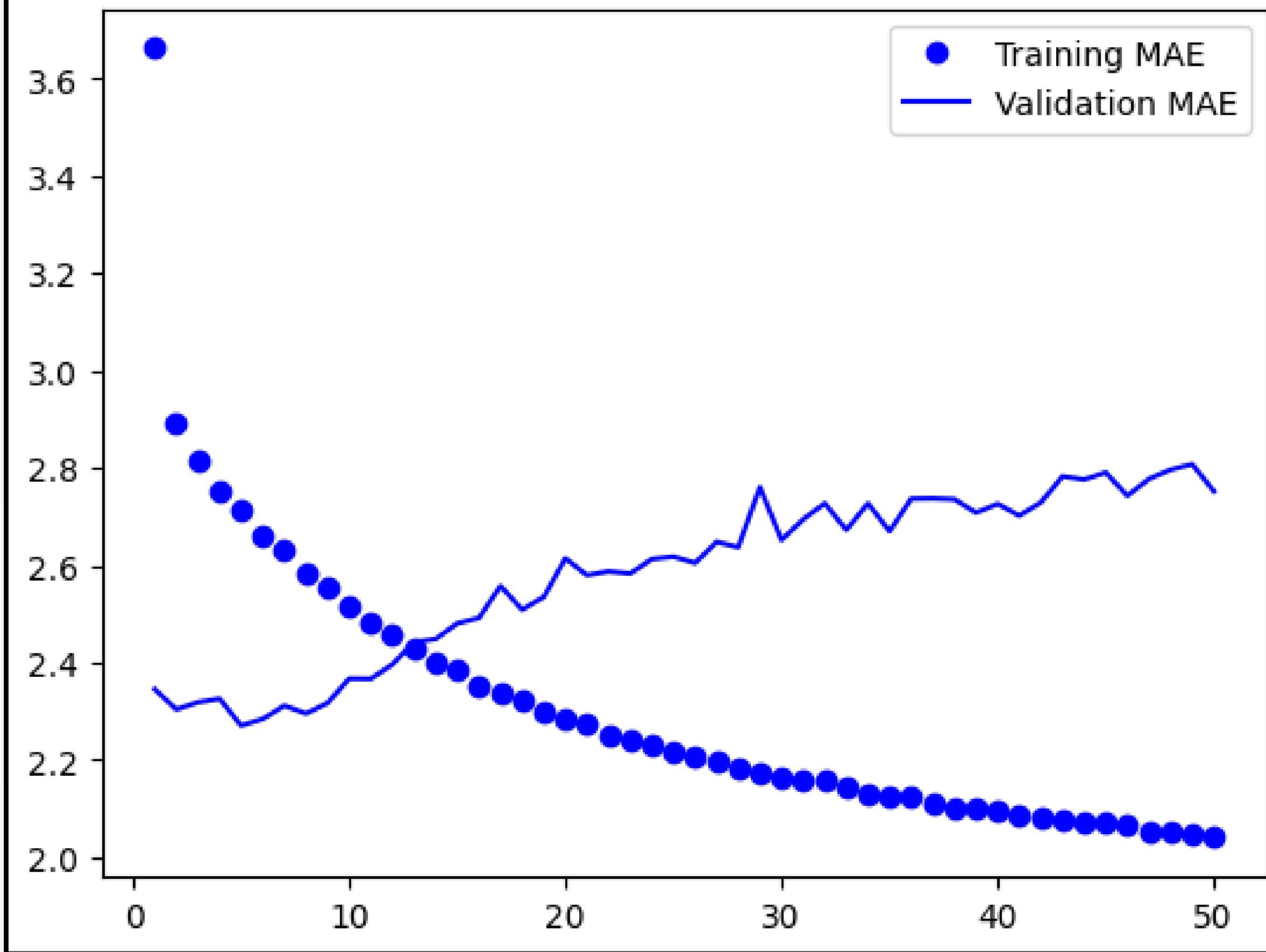
- 增加網路容量, 提升表徵能力
- MAE (2.39)
- Google 翻譯演算法就是由 7 個大型 LSTM 組成

Deep RNN

$$h', y = f_1(h, x) \quad b', c = f_2(b, y) \dots$$

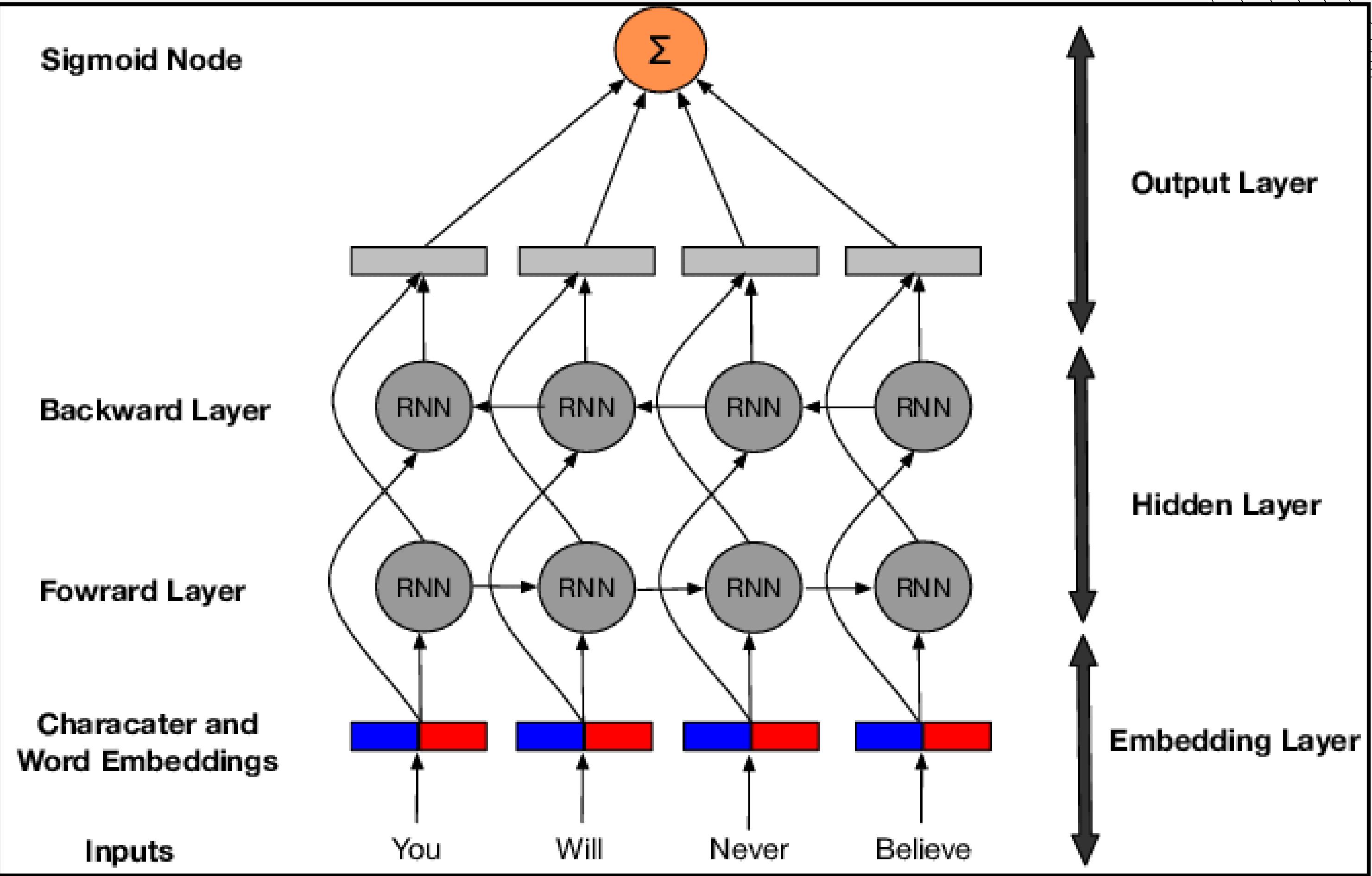


Training and validation MAE

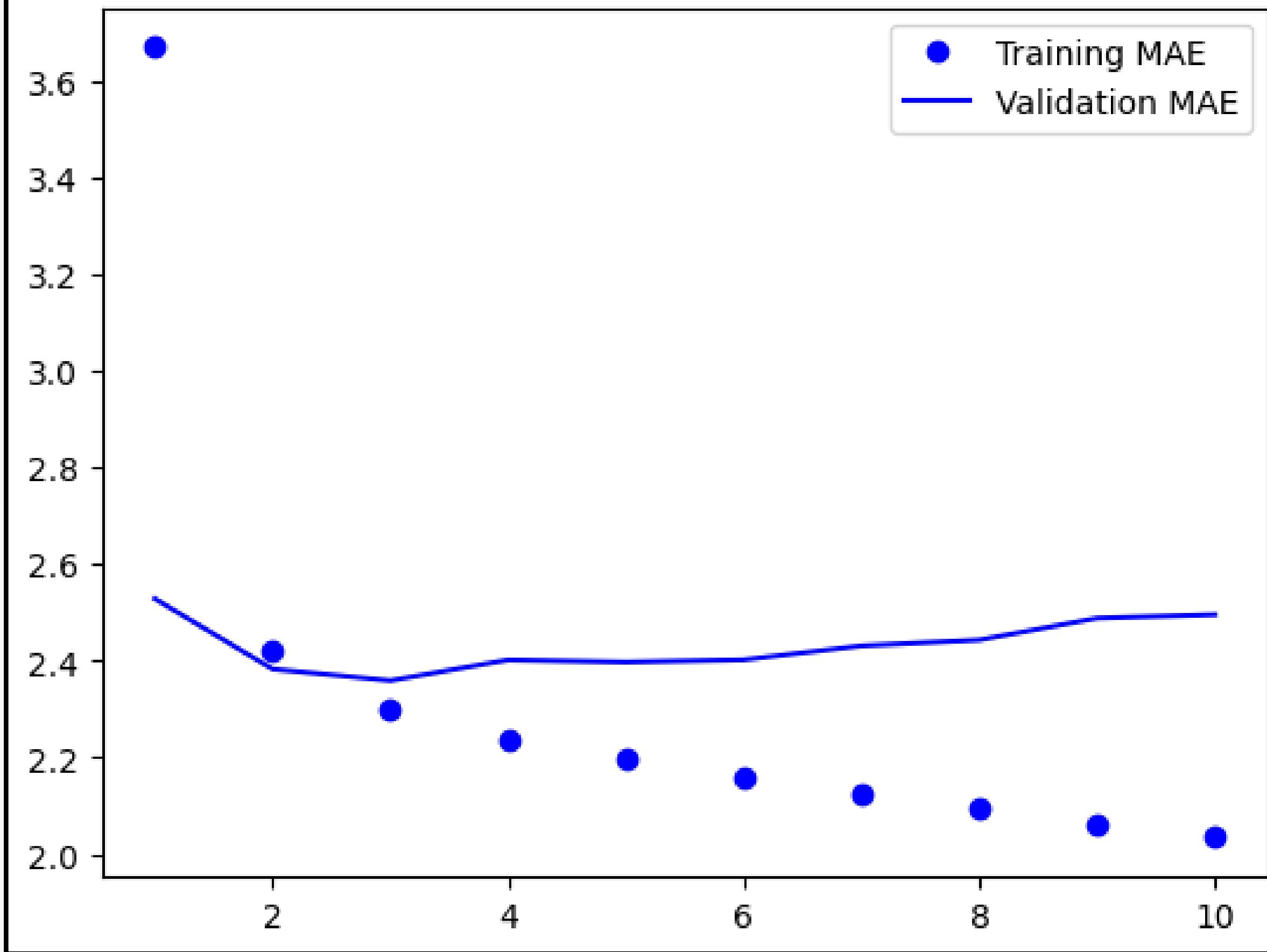


雙向 RNN

- 常用於自然語言處理
- 捕捉單向 RNN 可能忽略的 pattern
- 可使用正、反向序列, 捕捉更好的 pattern
- 因容量擴增, 導致 overfit 提早發生



Training and validation MAE



ENDING

