

目錄

Ch01 從感知器看神經網路的底層知識

- 1-1 最早的人工神經元 - Rosenblatt 感知器
- 1-2 增加感知器模型的能力
- 1-3 用線性代數實現神經網路模型

Ch02 梯度下降法與反向傳播

- 2-1 導數的基礎概念
- 2-2 以梯度下降法 (gradient descent) 對模型訓練問題求解
- 2-3 反向傳播 (back propagation)

Ch03 多層神經網路的建立與調校

- 3-1 動手實作：建立辨識手寫數字的多層神經網路
- 3-2 改善神經網路的訓練成效
- 3-3 實驗：調整神經網路與學習參數

Ch04 用卷積神經網路 (CNN) 進行圖片辨識

- 4-1 卷積神經網路 (CNN)
- 4-2 實作：以卷積神經網路做圖片分類
- 4-3 更深層的 CNN 與預訓練模型

Ch05 用循環神經網路 (RNN、LSTM...) 處理序列資料

- 5-1 RNN 的基本概念
- 5-2 RNN 範例：預測書店銷售額
- 5-3 LSTM (長短期記憶神經網路)
- 5-4 LSTM 範例：文字的 Auto-Complete 機制

Ch06 自然語言處理的重要前置工作：建立詞向量空間

- 6-1 詞向量空間的基本知識
- 6-2 做法(一)：在神經網路建模過程中「順便」生成詞向量空間
- 6-3 做法(二)：以 word2vec、GloVe 專用演算法生成詞向量空間

Ch07 用機器翻譯模型熟悉 seq2seq 架構

- 7-1 機器翻譯模型的基本知識
- 7-2 機器翻譯的範例實作
 - 7-2-1 tf.Keras 函數式 API 簡介
 - 7-2-2 建構模型前的工作
 - 7-2-3 建構模型
 - 7-2-4 訓練及測試模型
 - 7-2-5 實驗結果

Ch08 認識 attention 與 self-attention 機制

- 8-1 熟悉 attention 機制

- 8-2 認識 self-attention 機制
 - 8-2-1 self-attention 的基本概念
 - 8-2-2 self-attention 機制的算法
 - 8-2-3 multi-head (多頭) 的 self-attention 機制

Ch09 Transformer、GPT 及其他衍生模型架構

- 9-1 Transformer 架構
 - 9-1-1 編碼器端的架構
 - 9-1-2 解碼器端的架構
 - 9-1-3 Transformer 內的其他設計
 - 9-1-4 小編補充：觀摩 keras 官網上的 Transformer 範例
- 9-2 Transformer 架構的衍生模型：GPT、BERT
 - 9-2-1 認識 GPT 模型
 - 9-2-2 認識 BERT 模型
 - 9-2-3 其他從 Transformer 衍生出的模型

附錄 A 延伸學習 (一)：多模態、多任務...等模型建構相關主題

附錄 B 延伸學習 (二)：自動化模型架構搜尋

附錄 C 延伸學習 (三)：後續學習方向建議

附錄 D 使用 Google 的 Colab 雲端開發環境