

Day98

深度學習應用卷積神經網路

訓練卷積神經網路的細節與技巧 處理大量數據



出題教練

楊証琨



大數據？

- Cifar-10 資料集相對於常用到的影像來說是非常小，所以可以先把資料集全
部讀進記憶體裡面，要使用時直接從記憶體中存取，速度會相當快
- 但是如果我們要處理的資料集超過電腦記憶體的容量呢？桌上電腦的記憶體
多為 32, 64, 128 GB，當處理超大圖片、3D 影像或影片時，就可能遇到 Out
of Memory error

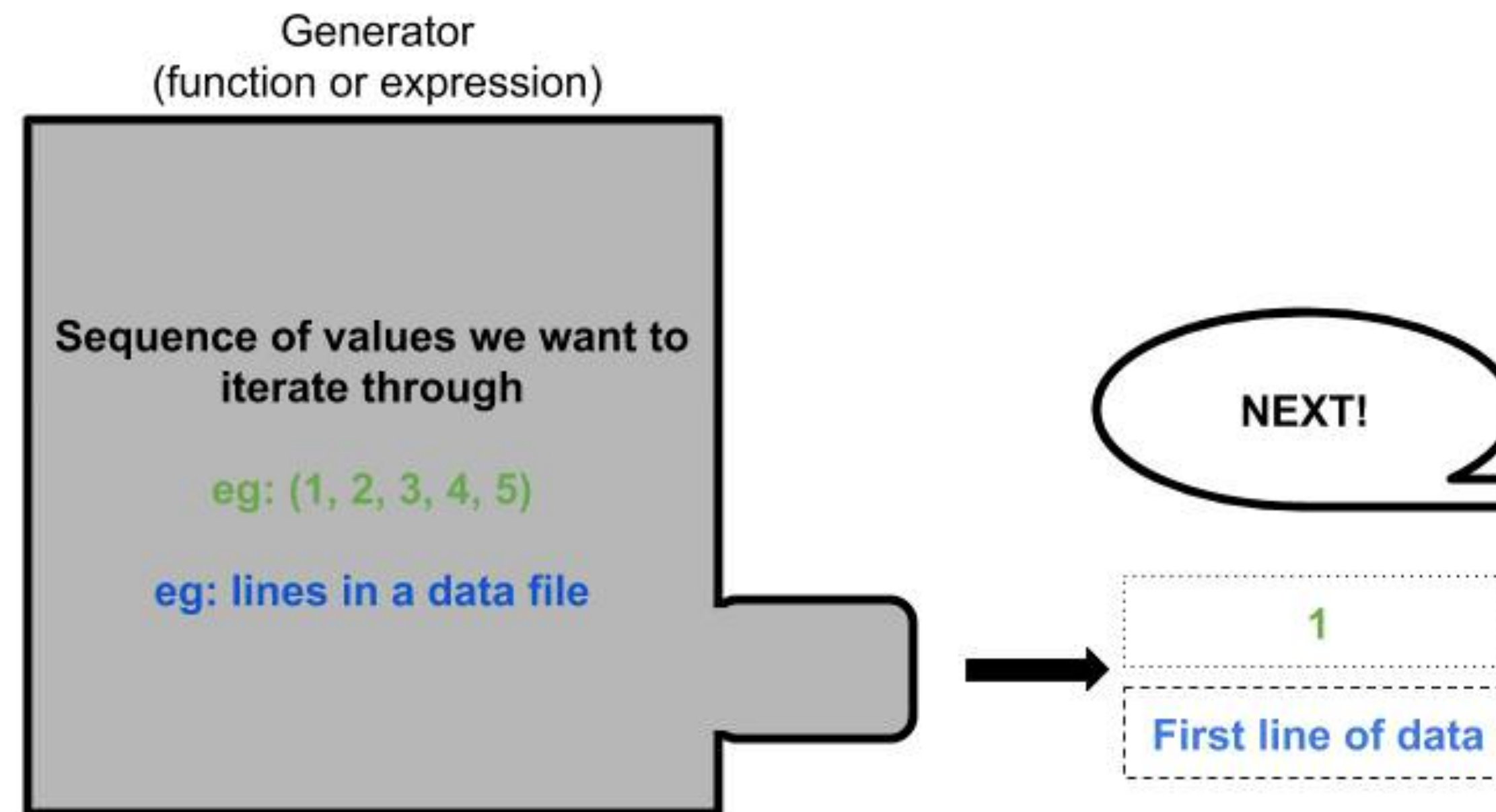
批次 (batch) 讀取

如同訓練神經網路時，Batch (批次) 的概念一樣。我們可以將資料一批一批的讀進記憶體，當從 GPU/CPU 訓練完後，將這批資料從記憶體釋出，在讀取下一批資料

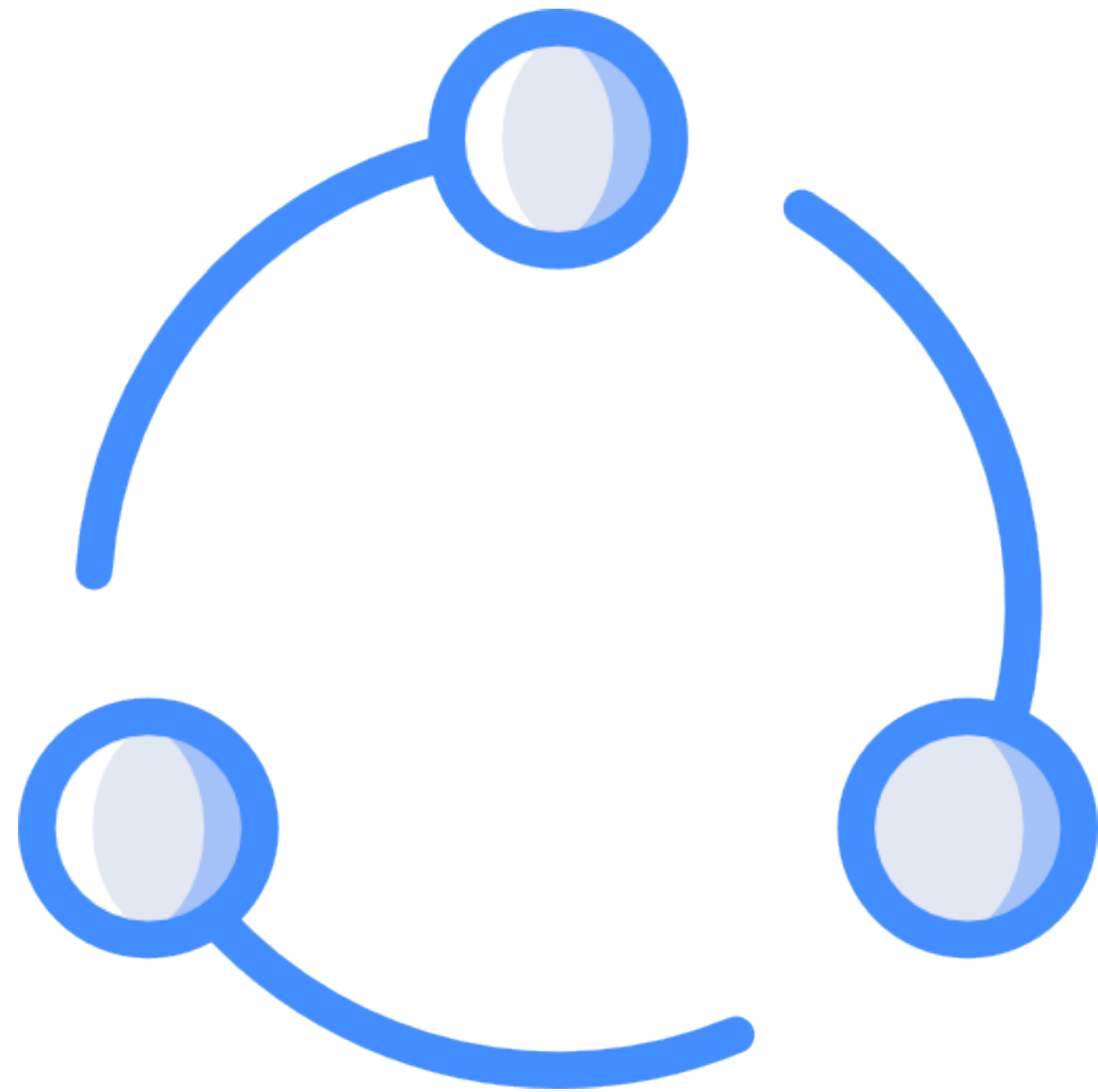


如何用 Python 撰寫批次讀取資料的程式碼

- 使用 Python 的 generator 來幫你完成這個任務！
- Generator 可以使用 `next(your_generator)` 來執行下一次循環
- 假設有一個 list，其中有 5 個數字，我們可以撰寫一個 generator，用 `next(generator)` 會自動吐出 list 的第一個數字，再用第二次 `next` 則會吐出第二個數字，以此類推



從程式碼學習！



- 將原本 Python function 中的 return 改為 yield，這樣 Python 就知道這是一個 Generator 囉
- 請參考本日程式碼，學習 generator 的寫法

解題時間 It's Your Turn

請跳出PDF至官網Sample Code & 作業
開始解題

