

D83：訓練神經網路的細節與技巧 - Batch normalization

[PDF 下載](#)[全螢幕](#)

Sample Code & 作業內容

請參考範例程式碼Day083_BatchNorm.ipynb，完成以下作業：

作業 1：試比較有 BN 在 Batch_size = 2, 16, 32, 128, 256 下的差異

作業 2：請嘗試將 BN 放在 Activation 之前，並比較訓練結果

作業 3：請於 BN 放在 Input Layer 後，並比較結果

作業請提交Day083_HW.ipynb

[檢視範例](#)

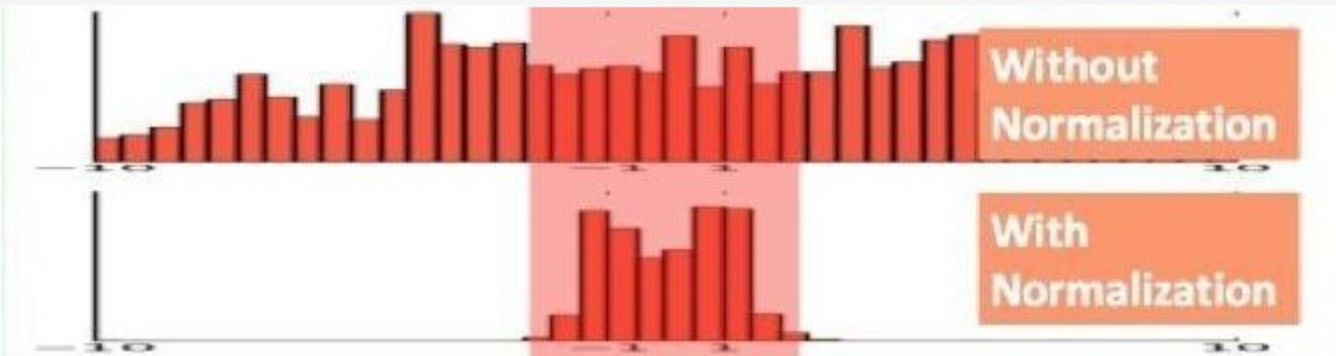
參考資料

知乎 - BatchNorm 的原理與實戰

- 背景：
 - 為什麼神經網路這麼不穩定 → 神經網路各層間的映射頻繁的變換 (Internal Covariate Shift)
- Internal Covariate Shift 的問題
 - 網路在初期需要透過不斷大幅調整來適應輸入的分布變化
 - 深層網路的梯度消失/爆炸問題使分布更加不穩
- 解法
 - 在各層網路中的各特徵軸 (channel) 單獨進行標準化，使其變為分布平均為 0, 變異數為 1 的分布，再加上平移與縮放，就可以表示原本的分布。

參考連結：

- [為何要使用 Batch Normalization – 莫煩 python](#)
- [Batch normalization 原理與實戰 – 知乎](#)



提交作業

請將你的作業上傳至 Github，並貼上該網址，完成作業提交

[確定提交](#)[如何提交](#)

到 Cupoy 問答社區提問，讓教練群回答你的疑難雜症

[向專家提問](#)[如何提問](#)