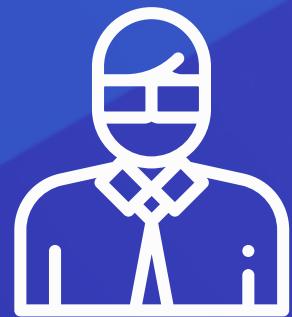




Day 65

深度學習理論與實作

深度學習體驗：啟動函數與正規化



出題教練

陳宇春 / 陳明佑



知識地圖 深度學習簡介

深度學習體驗 - 啟動函數與正規化

深度神經網路

Supervised Learning Deep Neural Network (DNN)

簡介 Introduction

套件介紹 Tools: Keras

組成概念 Concept

訓練技巧 Training Skill

應用案例 Application

卷積神經網路

Convolutional Neural Network (CNN)

簡介 introduction

套件練習 Practice with Keras

訓練技巧 Training Skill

電腦視覺 Computer Vision

深度學習簡介

Introduction of DNN

神經網路
歷史

深度學習
概念

深度學習
體驗

本日知識點目標

- 理解批次大小 (Batch size) 與學習速率 (Learnig Rate) 對學習結果的影響
- 經由實驗，體驗不同啟動函數的差異性
- 體驗正規化 (Regularization) 對學習結果的影響

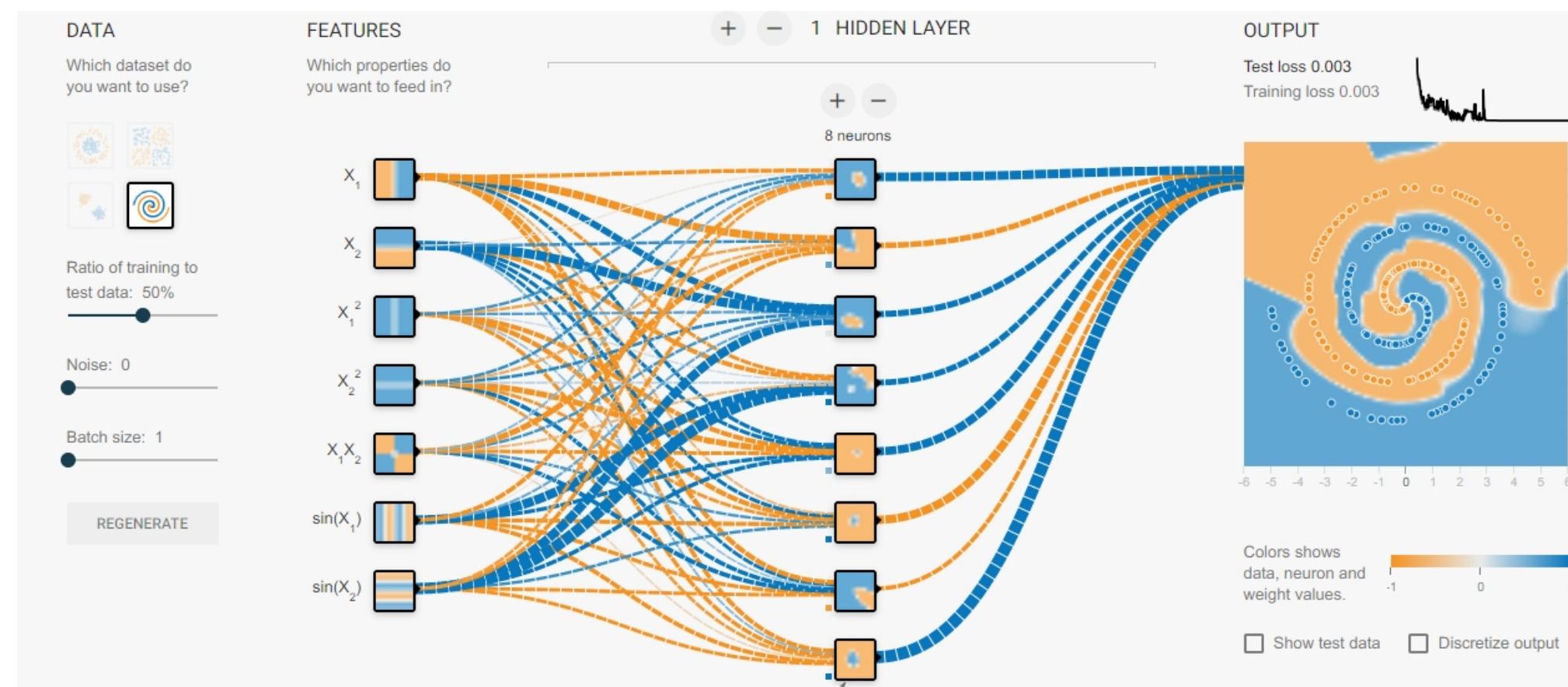
練習 5：切換批次大小

練習操作

- 資料集切換：分類資料集(右下)-螺旋雙臂，特徵全選，隱藏層 1 層 /8 神經元
- 調整 **不同的批次大小** 後執行 500 次遞迴，看看學習效果有何不同？

實驗結果

- 批次大小很小時，雖然收斂過程**非常不穩定**，但平均而言會收斂到較好的結果



**** 註：**實務上，批次大小如果極小，效果確實比較好，但計算時間會相當久，因此通常會依照時間需要而折衷

練習 6：切換學習速率

練習操作

- 資料集切換：分類資料集(右下)-螺旋雙臂，特徵全選，隱藏層 1 層 /8 神經元，批次大小固定 10
- 調整 **不同的學習速率** 後執行 500 次遞迴，看看學習效果有何不同？

實驗結果

- 小於 0.3 時 學習速率較大時，收斂過程會越不穩定，但會收斂到較好的結果
- 大於 1 時 因為過度不穩定而導致無法收斂



練習 7：切換啟動函數

練習操作

- 資料集切換：分類資料集(右下)-螺旋雙臂，特徵全選，隱藏層 1層 /8 神經元，批次大小固定 10，學習速率固定 1
- 調整 **不同的啟動函數** 後執行 500 次遞迴，看看學習效果有何不同？

實驗結果

- 在這種極端的情形下，Tanh 會無法收斂，Relu 很快就穩定在很糟糕的分類狀態，惟有 Sigmoid 還可以收斂到不錯的結果
- 但實務上，Sigmoid 需要**大量計算時間**，而 Relu 則相對快得很多，這也是需要取捨的，在本例中因位只有一層，所以狀況不太明顯



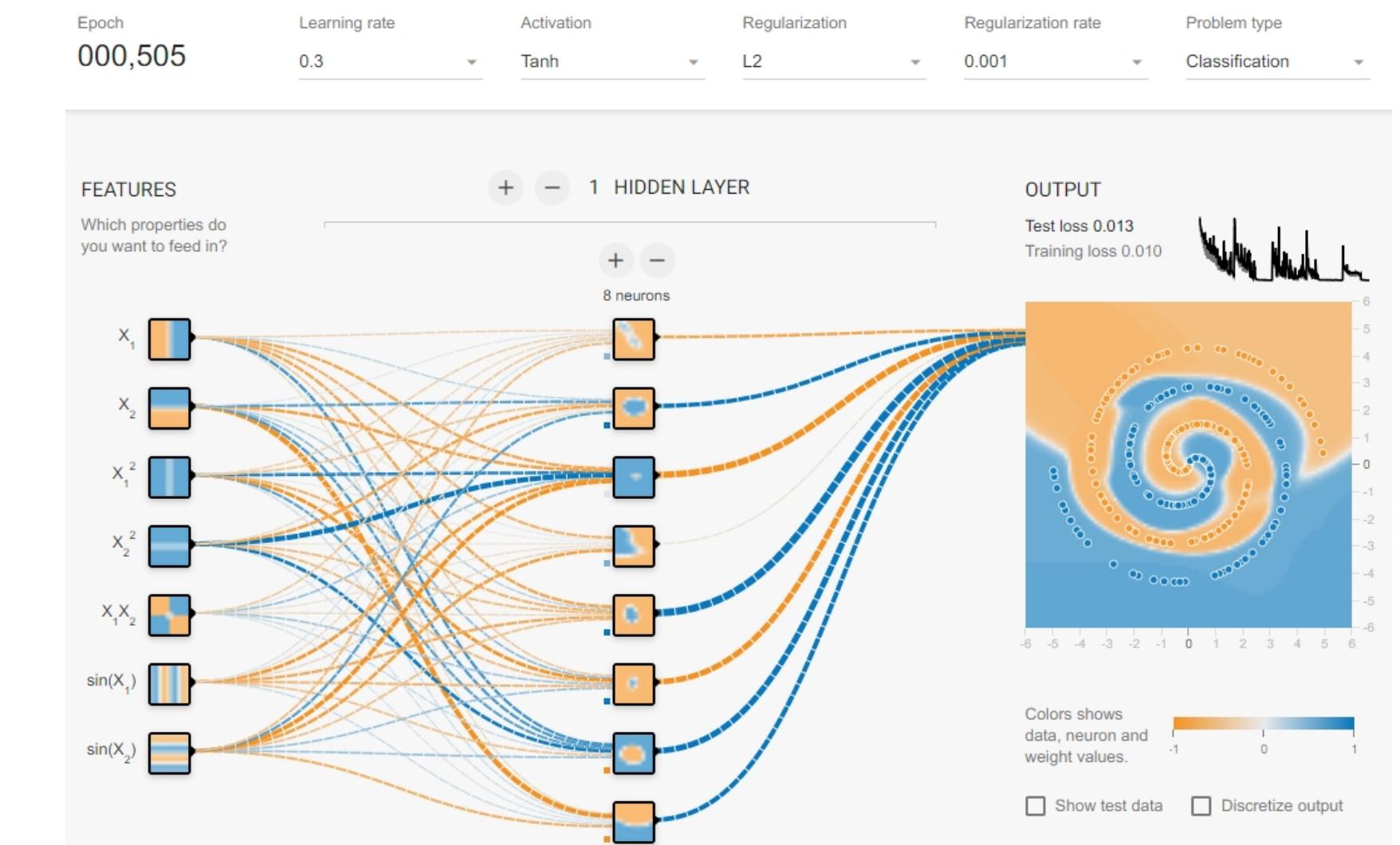
練習 8：切換正規化選項與參數

- 練習操作

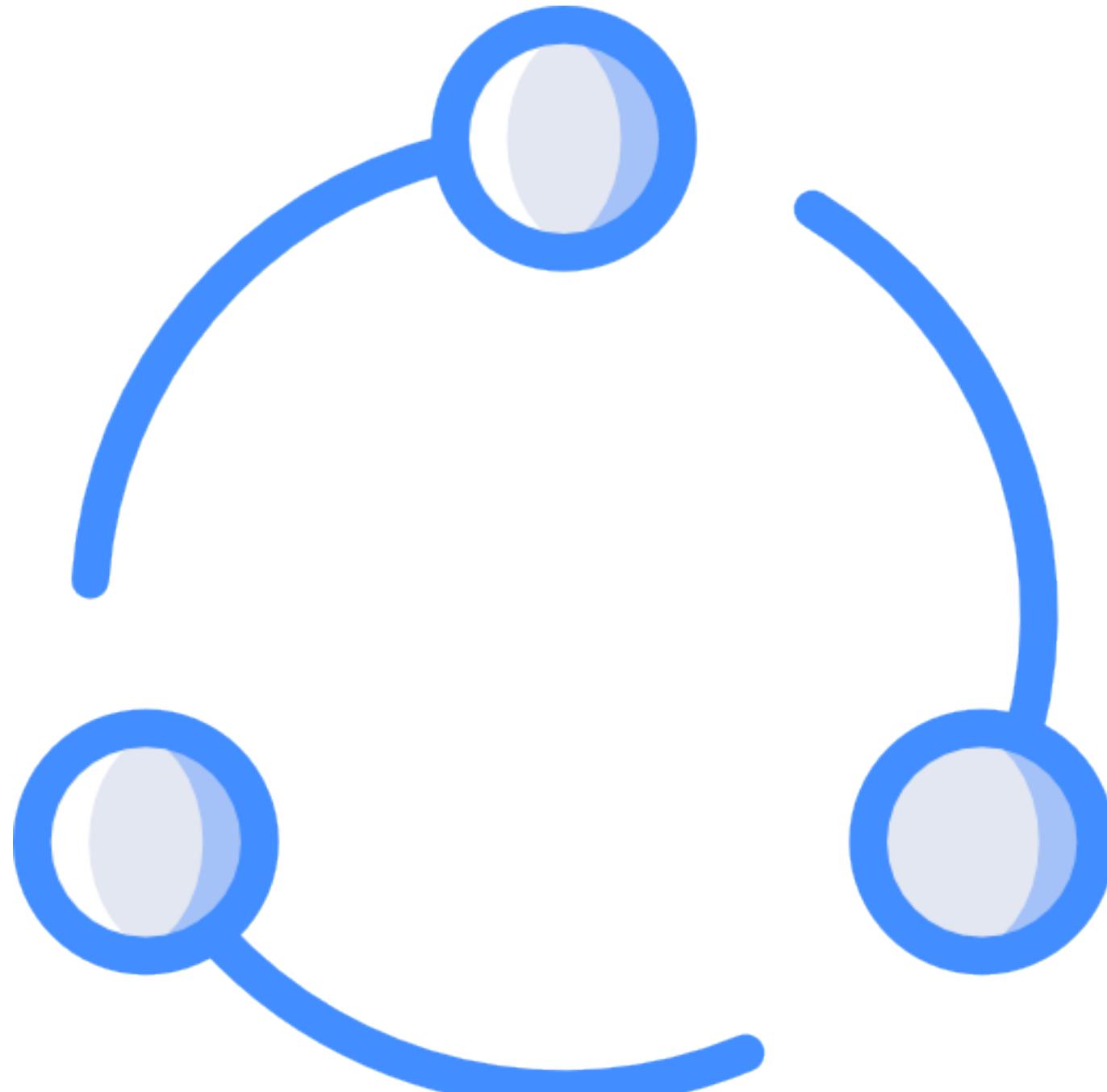
- 資料集切換：分類資料集(右下)-螺旋雙臂，特徵全選，隱藏層1層/8神經元，批次大小固定 10，學習速率固定 0.3，啟動函數設為 Tanh
- 調整不同的正規化選項與參數 後執行 500 次遞迴，看看學習效果有何不同？

- 實驗結果

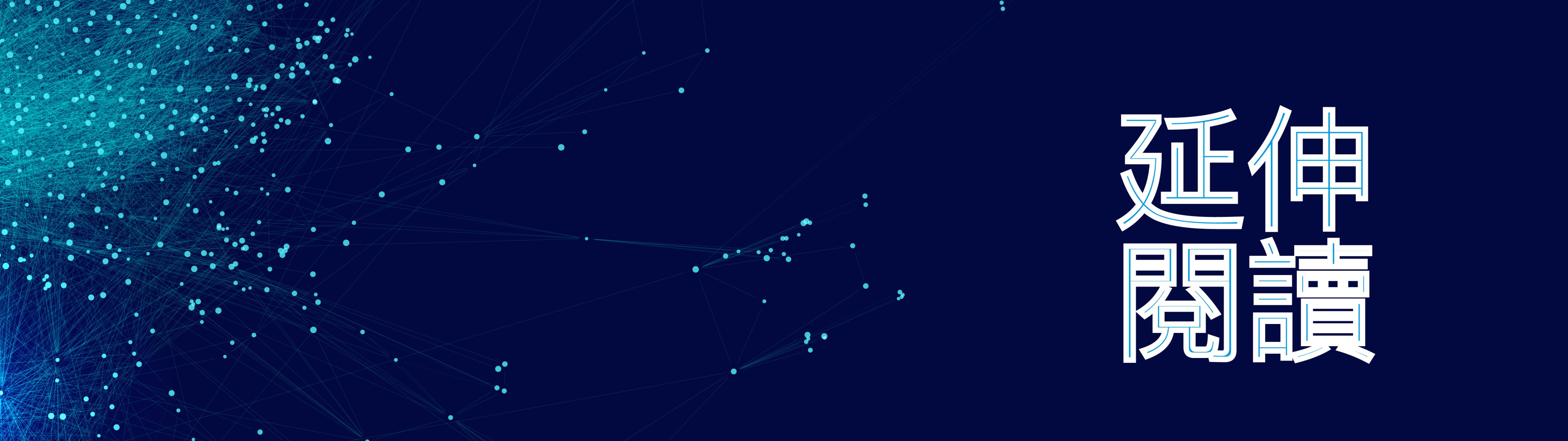
- 我們已經知道上述設定本來就會收斂，只是在較小的 L1 / L2 正規劃參數下收斂比較穩定一點
- 但正規化參數只要略大，反而會讓本來能收斂的設定變得無法收斂，這點 L1 比 L2 情況略嚴重，因此本例中最適合的正規化參數是 L2 + 參數 0.001
- 實務上：L1 / L2 較常使用在非深度學習上，深度學習上效果有限



重要知識點複習



- 批次大小**越小**：學習曲線越**不穩定**、但**收斂越快**
- 學習速率**越大**：學習曲線越**不穩定**、但**收斂越快**，但是與批次大小不同的是 - 學習速率大於一定以上時，有可能不穩定到**無法收斂**
- 當類神經網路層數**不多**時，啟動函數 Sigmoid / Tanh 的效果比 Relu 更好
- L1 / L2 正規化在非深度學習上效果較明顯，而**正規化參數較小才有效果**



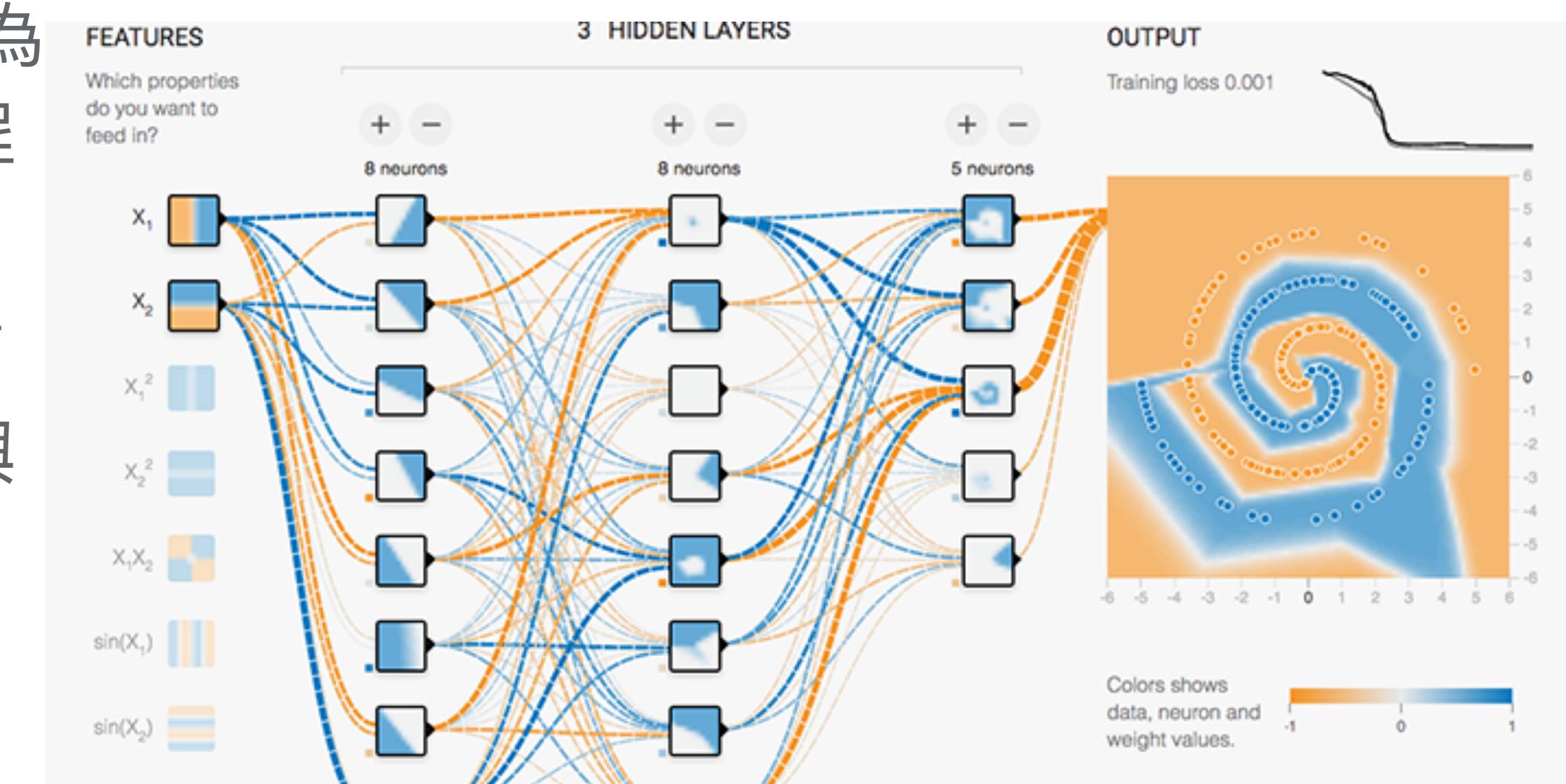
延伸 閱讀

除了每日知識點的基礎之外，推薦的延伸閱讀能補足學員們對該知識點的了解程度，建議您解完每日題目後，若有
多餘時間，可再補充延伸閱讀文章內容。

推薦延伸閱讀

Understanding neural networks with TensorFlow Playground
 Google Cloud 官方教學 [網頁連結](#)

- 這是 Google 官方出的 PlayGround 教學，除了與作業類似的跑參圖(但是沒有限制 epoch 500)，還有 Google 認為類神經網路初學者需要學習的內容，如果覺得我們的課程有不完整的地方，建議可以直接參考上面的說明
- 但還是老話一句：不要強迫自己現在就全部都看懂，細節的部分後面還有其他的課程會說明，先專注享受你最有興趣的部分吧。



推薦延伸閱讀

深度深度學習網路調參技巧 with TensorFlow Playground 知乎

作者：煉丹實驗室 [網頁連結](#)

- 如果你對於深度學習的參數調整有更多興趣，那麼建議你可以讀一讀這篇的內容
- 提到這篇，主要是讓大家看到「自動調參」的章節 (倒數第二段，總結之前)，可以知道深度學習的調參也不外乎是 GridSearch 與 RandomSearch，但是也與機器學習一樣，調整這些參數前最好能先了解它

深度学习网络调参技巧



萧瑟

深度学习 (Deep Learning) 话题的优秀回答者

802 人赞同了该文章

转载请注明: [炼丹实验室](#)

之前曾经写过一篇文章，讲了一些深度学习训练的技巧，其中包含了部分调参心得：[深度学习训练心得](#)。不过由于一般深度学习实验，相比普通机器学习任务，时间较长，因此调参技巧就显得尤为重要。同时个人实践中，又有一些新的调参心得，因此这里单独写一篇文章，谈一下自己对深度学习调参的理解，大家如果有其他技巧，也欢迎多多交流。

好的实验环境是成功的一半

由于深度学习实验超多，代码风格良好的实验环境，可以让你的人工或者自动调参更加省力，有以下几点可能需要注意：

- 将各个参数的设置部分集中在一起。如果参数的设置分布在代码的各个地方，那么修改的过程想必会非常痛苦。
- 可以输出模型的损失函数值以及训练集和验证集上的准确率。
- 可以考虑设计一个子程序，可以根据给定的参数，启动训练并监控和周期性保存评估结果。再由



解題時間

It's Your Turn

請跳出PDF至官網Sample Code & 作業
開始解題

