

# Lab 1 結報

姓名：林業語      學號：109611036

## 1. 請敘述在課堂上實作之神經網路的架構

在課堂上實作的神經網路架構為多層感知器（**Multilayer Perceptron, MLP**），由四層 **Dense** 層構成，**Activation function** 使用 **relu** 其中第一層有 512 個神經元、第二層有 256 個神經元、第三層有 128 個神經元、第四層有 64 個神經元，輸出層則有 10 個神經元，對應 10 個手寫數字的分類，**Activation function** 使用 **softmax** 輸出各類別的機率。整個模型的輸入為一維向量，需將原始手寫數字圖片先進行 **reshape**，將二維圖片資料轉換成一維向量，再進行訓練。

**Softmax** 是一種用於多類別分類問題的 **Activation function**。可以將一個向量映射到一個新的向量，新的向量的每個元素都表示為原始向量對應類別的機率。**Softmax** 函數的輸出值都在 0 到 1 之間，且所有輸出值的總和為 1，因此可以視為一種機率分佈。

這次 Lab 中 要判斷手寫數字 0~9 因此最終輸出層會是 10 個，也就是 0~9 的機率。

## 2. 請敘述 Loss Function 與 Optimizer 的作用

**Loss Function** 是用來計算模型輸出結果與實際 **Label** 之間的差異，也就是模型的預測結果與真實答案之間的距離，通常會將這個距離定義為一個 **loss function**。在這次 Lab 中，選擇的 **Loss Function** 為 **categorical\_crossentropy**，這是一種常用於多分類問題的 **loss function**。

根據我在網路上找到的資料，**categorical\_crossentropy** 的優點是可以處理多類別分類問題，而且可以直接將概率分布作為損失函數的輸入，不需要進行任何轉換。它還具有比較好的數學性質，能夠避免過度擬合，因此在實際應用中經常被使用。

另外，**categorical\_crossentropy** 對於類別標籤是 **one-hot** 編碼的情況比較適用，如果類別標籤不是 **one-hot** 編碼，而是整數編碼的情況，就需要使用 **SparseCategoricalCrossentropy** 損失函數。

**Optimizer** 是用來更新模型權重的方法，通過優化算法來調整模型參數，以最小化損失函數的值，從而提高模型的準確度和效能。

在深度學習中，通常使用反向傳播算法（**backpropagation**）計算梯度，然後使用梯度下降算法（**gradient descent**）更新權重。梯度下降算法有多種變種，其中一種是 **AdamOptimizer**。**AdamOptimizer** 可以自動調整學習率，適應性地調整參數更新的速度。**Adam** 優化器的主要優點是快速收斂和對參數不同尺度的適應性調整。

### 3. 心得

在這個 Lab 中，我們使用 Tensorflow 的 Keras 建立了一個深度學習模型，並利用 MNIST 手寫數字資料集進行訓練和測試。這個 Lab 讓我更深入地了解神經網路的基礎概念和常用技巧，例如資料前處理、建立模型、訓練與評估等。此外，我們也使用了 matplotlib 來繪製圖表，以更好地呈現訓練過程和結果。這個 Lab 是深度學習入門的實踐，整體還算容易完成，我也期待能在未來的學習中深入探索更多深度學習相關的應用和技術。

參考資料：

<https://zhuanlan.zhihu.com/p/105722023>

[https://www.tensorflow.org/api\\_docs/python/tf/keras/losses/CategoricalCrossentropy](https://www.tensorflow.org/api_docs/python/tf/keras/losses/CategoricalCrossentropy)