**Lab 1 結報**

姓名：賴昱凱 學號：111511141

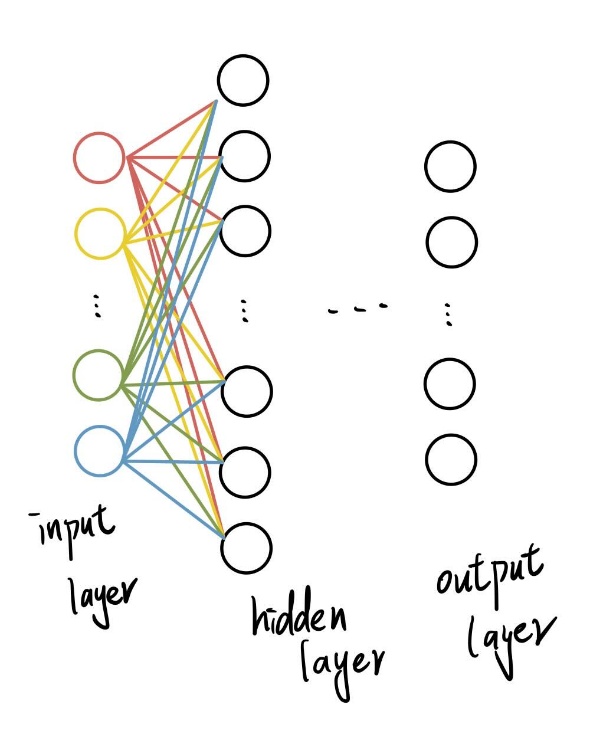
1. 請敘述在課堂上實作之神經網路的架構

一張含有 文字, 螢幕擷取畫面, 字型 的圖片

AI 產生的內容可能不正確。

本模型是由4層hidden layer以及input、output layer，共6層所組成的DNN模型，DNN模型的全稱為Deep Neural Network，顧名思義代表所有的節點都與前後層的所有節點相連，利用權重做相乘往下一層前進，因此整個每層共有個參數，以本模型為例：

全部相加即為本模型的總參數：575050



除了linear以外，還有一個部分稱為activation function，他是非線性的layer，因為若整個model都是linear的layer，會很難處理現實中的問題，因為現實中的問題基本上都不會是linear的，因此需要在model中加上activation function來讓模型更加靈活。

本模型皆使用ReLU這個最常見的activation function，他會將正數原封不動的送給下一層，而遇到負數就將其設為0，這是一種最不耗計算資源的non-linear方式。下圖還有更多常見的activation function。

一張含有 文字, 圖表, 字型, 數字 的圖片

AI 產生的內容可能不正確。

本模型的輸出層為有10個節點的神經層，最後一層利用softmax作為actiavtion function，利用指數化確保輸出皆為正數，再讓所有輸出的值都介於0~1且相加為1，代表各種類的可能性，最高的那個就代表此次模型預測的結果，這是幾乎所有分類問題處理輸出最基本的方式。

Softmax:

1. 請敘述Loss Function與Optimizer的作用

Loss Function:

衡量模型預測值與真實值之間的差距，通常來說loss function的目標是值越小越好，代表與真實值的差距越小。常用的loss function有：

Mean Square Error

由方程式可知，他是直接計算模型輸出與真實值的距離，因此非常適合用於回歸問題，例如氣溫變化等。

Cross Entropy

通常會是一個N維的one-hot編碼（N代表分類數量），也就是只有一個值為1，其他皆為0，這個1就代表data的分類；而會是一個機率分佈，代表模型認為此data為各種類的機率為何。

根據方程式我們可以發現，當預測答案為正解的機率很高時，式子所算出的loss會很低，例如，而其他值都會是0；反之當正解被預測得很低時，式子將會算出十分高的loss，例如。因此Cross Entropy十分適合應用在分類問題上。

Optimizer:

Optimizer的作用是根據Loss Function的結果來調整模型的參數，讓模型預測得越來越好，最終目的是找到一組參數使loss最小化。可以把Loss想像成地形圖，Optimizer的功能就是設計一種算法，根據地形圖的斜率（梯度）讓loss往山谷走，達到地形的最低點。

常見的Optimizer有SGD、Adam等，SGD就是單純的觀察梯度進行參數的更新，而Adam等其他更複雜的Optimizer還會根據動量或其他因素動態更新參數，讓模型訓練收斂更快更穩定。

1. 心得

本次的實驗是module 2的第一次課程，先簡單的講解無線通訊以及人工智慧最基本的知識，讓我們了解課程圍繞的主題。而本次的實驗內容是很基礎的手寫數字分類，讓我們熟悉環境以及model training相關程式語法的使用。基本上我們都很快的完成作業，希望可以在未來的課程中順利完成並學習到AI在無線通訊的應用。