神秘的典禮

**壹、前言**

這一天，穿著一襲禮服的Win10參加了Microsoft學校的校慶典禮，當大家高興地聊著天吃著東西時，突然一個愚笨的少年衝了過來，用兇狠的眼神瞪著她，他自稱他原本是Microsoft學校裡的一個工友，只是因為他太愛國了，屢屢遭到學生投訴，今天他來就是要來叫Win10大美女幫他討回公道。

他說:「反正同學們對我的成見已深，我即使回到了學校也是不太可能好好工作的，妳如果沒辦法幫助我找到第二份工作，我就會剝奪你最珍惜的東西，然後這間學校也會付之一炬。」這時，所有在典禮的同學們呆掉了，這是他們第一次這麼靠近死亡，這麼希望能夠離開這間學校，大家的性命都託付在Win10的手裡，Win10緊張的說不出話來，那個少年說話了:「不要再拖了!時間一直在流逝，妳快快想辦法，我只給妳一個禮拜的時間，如果沒幫助我找到第二份工作，後果不堪設想!」Win10靠著她的智慧與勇氣，想出了一個方法，「以人工智慧協助**求職者面試**預測」，因此她號召了全校最優秀且頂尖的科學、資訊人才，開始著手訓練這個模型。

經過她們不眠不休的測試與訓練，事情似乎露出了一絲曙光，Win10親自教導同學們如何計算矩陣及向量，同學們一聽了她的講解，恍然大悟，也讓整個模型開發越來越順利，但，她們卡關了，他們卡在一個十分重要的點”激勵函數”，沒有了激勵函數，深度學習僅僅是數個矩陣進行運算，只有經過非線性的激勵函數，才能真正讓模型有意義，同樣也是資訊能力超強的你們，加入她們的行列吧!幫助思維謹慎、具有領導能力的Win10守住校園吧，校園是否會被這個愚笨的不良少年摧毀，全靠你們的智慧了。

類神經網路中都需要一個激勵函數(Activation Function)，激勵函數的主要用途是為了解決輸入和輸出的線性關係、並增加模型的表現力，因此有許多不同的激勵函數，這些激勵函數各自有各自的優缺點，而其中以ReLU函數成為現代深度學習的主流，其他的函數有許多問題，像是Sigmoid和Tanh函數時常會發生梯度消失的問題，再者，ReLU相較其他函數而言，十分容易計算，也可以省下許多資源。

深度學習時常用到的目標函數(Object Function)都是使用損失函數(Loss Function)，也就是下面題目要叫你算的Loss，一個模型的好壞通常與損失函數的設計也有很大的關係，回歸函數被分類為兩個部分，回歸及分類，對於回歸問題，Loss Function 常用的是MSE均方誤差、MAE平均絕對值誤差這兩種，像是第59屆科展電腦與資訊學科的第一名「**以類神經網路為輔助自動生成小提琴演奏骨架**」就是使用MSE作為Loss Function，而在分類問題上，選擇的通常就是cross-entropy，當然，與激活函數一樣，每一個函數都有自己的優缺點，在開發時就是要對於模型多測試，找出最適合模型的函數。

參考下圖Image1可以發現常見的三種激活函數的圖形。此題只需用到第三個ReLU函數，我們可以推導出他的一般式:

OR

再來，請參考Image2的例題圖片及Image3的矩陣乘法範例圖片，本題由於不會使用到陣列，因此會固定向量、矩陣的大小，且你只需要幫忙求出Loss函數(這裡使用的是MSE均方誤差)的值即可，不需要進行偏微分(我真的人超級好的，請感謝Win10)。

附圖:

一張含有 文字, 地圖 的圖片

自動產生的描述

Image1 、 本圖出自: 交大人才培育計畫 Week4講義

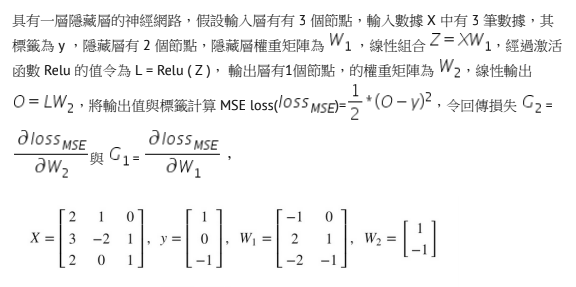


Image2 、 本圖出自: 交大人才培育計畫 期末測驗

一張含有 文字, 白板 的圖片

自動產生的描述

Image3 、 矩陣乘法的運算

貳**、輸入說明**

本題假設為有一隱藏層的神經網路，輸入層有3個節點，輸入數據X (一) 有3筆數據，標籤為Y (二) ，隱藏層有2個節點，隱藏層權重矩陣為 (三) ，線性組合，經過激活函數ReLU，令，輸出層有一個節點，其權重矩陣為 (四) ，線性輸出，計算。

(題目與上面附圖大致相同)

一、第一行會輸入9個數字，為矩陣X，格式請參考下面解說。

輸入 a b c d e f g h i

其意涵為:

二、第二行會輸入3個數字，為矩陣Y，格式請參考下面解說。

輸入 j k l

其意涵為:

三、第三行會輸入6個數字，為矩陣，格式請參考下面解說。

輸入 m n o p q r

其意涵為:

四、第四行會輸入2個數字，為矩陣，格式請參考下面解說。

輸入 s t

其意涵為:

參**、輸出說明**

計算並輸出Loss函數的值。

輸出到小數點後第二位。

行尾換行

肆**、範例測資**

輸入:

2 1 0 3 -2 1 2 0 1

1 0 -1

-1 0 2 1 -2 -1

1 -1

輸出:

22