機器學習專案作業一

本作業主要目的為使用Python語言及Keras實作類神經網路進行預測類別及數值預測。

分組：每組最多四人，請自行選出一人擔任組長。請將小組成員資料寄給助教郭哲宇。

Anaconda：https://www.anaconda.com/download/

作業內容包括：

* + 安裝與練習使用Anaconda裡的Jupyter Notebook系統。有問題可請教助教。
  + 利用Python及Keras建構兩個類神經網路分別進行Adult dataset的類別預測及數值預測。
    - UCI ML repository Adult dataset: <https://archive.ics.uci.edu/ml/index.php>
    - 數值預測方面，預測 hours-per-week；預測模型的績效衡量指標為RMSE
    - 類別預測方面，>50K or ≤50K；預測模型的績效衡量指標為Accuracy
    - 說明進行何種資料前置處理？如何處理缺失值？如何處理種類型屬性值？
    - 說明類神經網路的架構：隱藏層、各層節點數、激活函數、優化器、Regularization
    - 訓練資料(adult.data)、測試資料(adult.test)。分別列出對訓練資料及測試資料的預測績效
  + 專案程式直接上傳GitLab ( <http://140.125.84.84/> )，專案名稱：hw1+組長學號，並新增助教為專案成員，以利下載評分。
  + 請遵循學術論文寫作格式撰寫專案報告，內容至少必須包括下列要項：

**摘要** (學術論文摘要通常不分段，內容簡要地涵蓋動機、目的、方法與實驗結果)

**一、緒論**

**1.1動機** (探勘所選用的資料集之動機)

**1.2目的** (想進行哪種探勘，期望了解甚麼？)

**二、方法**

(1. 簡單說明程式架構，2. 執行程式的方法，無法執行程式，視為不正確)

**三、實驗**

**3.1資料集** (簡介資料集：名稱、筆數、欄位、型態…、並以資料表格呈現部份資料內容)

**3.2前置處理** (例如，欄位值正規化至0~1之間等。另外，須說明種類型屬性值得處理方法)

**3.3實驗設計** (實驗如何進行、若有參數的話，參數如何設定等）

**3.4實驗結果** (呈現結果)

**四、結論**

* 評分標準

內容(難度、豐富性、方法正確性)：60%，論文格式：20%，寫作技巧(文筆通順)：20%

* **任何複製剪貼網站資料或學長姐以前作業，皆屬抄襲，該作業分數最多不超過50分，嚴重者零分。**
* 此份報告最多不超過15頁，請加註頁碼。請遵循投稿論文寫作格式，使用一欄式，無需製作論文目錄。可以觀摩 ACM Transactions on Knowledge Discovery from Data 期刊論文的寫作技巧與格式。
* 紙本於上課前繳交或直接放入許老師的系辦信箱。電子檔(轉成PDF格式)繳交到網路學園。缺紙本或電子檔都視為未完成繳交作業。電子檔命名：HW1-組長學號。
* 繳交紙本時，在左上角訂一訂書針即可，無須其他任何裝訂。
* 繳交日期：109/04/16（四）。每遲交一天，扣該次作業成績10分。