資料探勘 專案二

* 本作業主要目的為分類器之訓練與分類預測。
* 資料探勘資源：<http://www.kdnuggets.com/>
* 作業內容包括：
  + 分類技術：分別使用SVM及Adaboost，並比較預測類別之績效。
  + 資料集 (ML Repository <https://archive.ics.uci.edu/ml/index.php>)
    - 請從ML Repository中自選兩個資料集，第一個為純數值型態資料集，第二個為混合型態資料集，也就是包含有數值型及非數值型態屬性 。
    - 需和第一次作業不相同的資料集。每個資料集資料至少一千筆以上。
    - 分別隨機將資料集分成：訓練資料80%，測試資料20%。
    - 須說明所選用的績效衡量指標。
  + 專案程式直接上傳GitLab ( <http://140.125.84.84/> )，專案名稱：hw2+組長學號，並新增助教為專案成員，以利下載評分。
  + 撰寫報告，請遵循學術論文寫作格式，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. 簡單說明所選用的探勘演算法或所採用的方法，2. 程式操作方法，無法操作者，以不正確論。)

**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 (**如果有的話，例如缺值與雜訊處理、值域正規化、屬性選擇或轉換、…**)**

**3.3實驗設計** (實驗如何進行、參數如何設定、績效如何評估等）

**3.4實驗結果** (呈現結果、附上一些畫面，且分析比較不同參數的預測績效)

**四、結論**

**五、參考文獻**

* 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，嚴重者零分。引用網站或他人資料，應註明引用來源。**
* 此份報告最多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：HW2-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。
* 每遲交一天，扣該次作業成績10分。
* 繳交日期：108/11/21（四）