

112-2 資料探勘期中報告

各位修習資料探勘的同學們好，以下為該學期的資料探勘期中報告的相關內容，以組別為單位，從下列兩種方式擇一進行報告：程式實作或論文閱讀報告，報告時間為期中考週，屆時請各組準備能清楚呈現組別成果之投影片上台報告。

1. 程式實作：

不限程式語言實作一個 KNN (不可使用 scikit-learn 或其他套件)，以及一個其他資料分類演算法 (例如: Random Forest, 可使用 scikit-learn 或其他套件)。規定資料集如下連結，內容為病人之各項測量值(如 BMI、年齡等)與該病人是否罹患糖尿病。請將程式流程圖、兩個實驗結果(實驗 A 及實驗 B)及分析上傳至網路大學供作為評分的參考。

資料集連結：https://drive.google.com/drive/folders/1T51gIXWLrd-VrS-bwqnHqxf7G1BEho1W?usp=share_link

2. 論文閱讀報告：

挑選一篇與分類或資料探勘相關論文上台報告。決定好要報告的論文後請先寄信給老師告知「組別、論文標題、論文網址」，確認該論文是否合適，並在網路大學公告的網址中留言，論文不可重複，先搶先贏。

◆ 需繳交資料為：

1. 書面報告
2. 投影片
3. 程式(論文報告組別不用繳交)

※ 每組只需一個組員上傳至網路大學，不用每個人都上傳。

◆ 注意事項：

- (1). 程式與論文閱讀均需上台報告，依照組別順序來報告。請於上課前提前來將投影片或程式放入報告的電腦內以節省時間。報告時間約 8~10 分鐘(報告 8 分鐘、問答 2 分鐘)，**請務必控制報告時間，不可超過**。經比對系統確認為抄襲，該次作業 0 分。
- (2). 擇程式報告或論文報告均需在報告當天製作投影片並繳交書面報告，書面報告依照給定格式撰寫，程式報告組別書面至少 800 字以上，論文報告組別書面至少 1500 字以上。
- (3). 書面報告及上台報告分數各佔 50%。
- (4). 實現的演算法越多個或閱讀越多篇報告，成績越佳。
- (5). 書面報告請採用黑白及雙面列印，調整為適合列印且清楚的格式，請不要直接將黑色為底的程式碼直接截圖放在報告。