

# 資料探勘 (Data Mining)

周永振  
**Yung-Chen Chou**

Department of Computer Science and Information  
Engineering, Asia University

E-mail: [yungchen@gmail.com](mailto:yungchen@gmail.com)

Web: <http://120.108.116.237/~jackjow>

# 大綱

- 資料探勘是一種能夠在**大量資料**中**自動化**發現一些**有用資訊**的流程，它可用來處理大量的資料，而且可以自其中發現一些之前未知的新奇、有價值的樣式。

# 學習目標

- 學習如何從眾多的資料中，以資料探勘（Data Mining）的技術，取得有用的資訊並進行資料的歸納與分析
- 要有能力以程式來實現本課程所教授的各種方法

# 教材

- 上課用書：
  - **資料探勘** (Introduction to Data mining), 施雅月、賴錦慧譯, 歐亞書局有限公司, ISBN: 978-9860154-657-5
  - Pang-Ning Tan, Michael Steinbach, Vipin Kumar, **Introduction to Data Mining**, Pearson Education Asia LTD, 2006.
- 參考書
  - **資料探礦概念與技術** (Data Mining: Concepts and Techniques, 2/e), 曾龍譯, 維科圖書出版, ISBN: 957-8675-92-5
  - **資料探勘** (Data Mining: A Tutorial-Based Primer), Richard J. Roiger, Michael W. Geatz, 東華書局 (ISBN: 986-7594-01-0)

# 課程設計

- 分爲“**課程講授**”和“**期末報告**”兩部分。
- **理論講授**：
  - 課堂講授Data Mining的相關理論 (約3/4學期) ，
- **期末報告**：(三選一)
  - 請於**期中考後兩週內**將組別與題目彙總繳交 (**逾期不候**) 。
  - **畫面報告**：由同學分組 (1~2人一組) ，自行尋找Data Mining相關議題進行書面報告撰寫，並於期末安排課堂報告。(60~80分)
  - **程式撰寫**：以2人爲一組，自行尋找Data Mining相關程式題目來實作與報告撰寫，並於期末做課堂報告與展示。(70~95分)
  - **應用系統模擬**：以2人爲一組，自行尋找Data Mining相關現成的應用軟體來實作一個議題與報告撰寫，並於期末做課堂報告與展示。(65~95分)

# 期末報告題目

- 書面報告題目訂定：
  - 國內外碩博士論文: (中文) 一人份, (英文) 二人份
    - 國家圖書館
    - 程式撰寫與應用系統模擬亦可由此找尋題目
  - 找尋會議論文(一人份)
    - 國內中文會議二篇
    - 國外英文會議一篇
  - 找尋期刊論文
    - 國內中文期刊(一人份)
    - 國外英文期刊(二人份)
  - **同學找到文章後須先給老師過目，若找不到文章的組別，統一由老師指定文章。**
  - 國家圖書館、本校圖書館、Google Scholar、Yahoo學術搜尋引擎、CiteSeer (<http://citeseer.csail.mit.edu/>)

# 評分標準

- 期中考：30%
- 期末報告：40%
  - 期末報告繳交頁數下限為10頁，(1.5倍行高，內文字型大小12點，標題文字16點為原則)，上限20頁。**各組於口頭報告日當天繳交。**
  - **每組口頭報告？分鐘，提問 5 分鐘。**
  - 禁止直接Copy他人文章 (以超過6成的報告內容與他人文章內容的重疊性為主)，違者**報告成績0分**。
- 平時成績：30%
  - 小考：20%
  - 課堂表現：10%
- 加分：程式撰寫，約5~10% **(學期末前兩週停收)**

# 電子郵件撰寫範例

Dear XX 老師,

信件內容...

Sincerely yours,

學生 王大華