

主導課程一：資料探勘與應用 (Data Mining : Concepts, Techniques, and Applications)

課程基本資料

開設學校：清華大學

開授教師：陳宜欣老師

班級人數：5000人 (保留250人給清大，聯盟校自訂，若為封閉型授權盟校，不超過100人)

開課級別：研究所開放大三 (含) 以上選課

授課語言：英文

授權方式：混成式

若選擇混成式，請附註混成式條件：協同教師須經主導教師同意，條件式授權盟校：成績考核方式由協同教師自訂。

協同教師學經歷建議：資工電機背景，過去研究和資料探勘相關

同步遠距上課時間：星期一 9:00~12:00

是否接受非同步授課：是

實體期末評量時間：2025/12/8

遠距上課位置：

https://www.youtube.com/@NTHU_ISA5810_DataMining

課程網頁：<https://www.cs.nthu.edu.tw/~yishin/courses/ISA5810/ISA5810->

聯盟學校修課人數與助教比例：每__50__名學生需__1__名助教

課程概述

Data mining serves as a crucial field that leverages advanced algorithms to reveal hidden, yet invaluable insights buried within extensive datasets. These algorithms are drawn from a multitude of areas such as machine learning, artificial intelligence, pattern recognition, statistics, and database systems, working together to facilitate a deeper understanding and analysis of data.

This course is designed to equip you with the foundational knowledge and hands-on experience needed to delve into the expansive world of data mining. Whether you are looking to enhance your skill set or embark on a new career path, this course will serve as a stepping stone to achieving your goals. The curriculum encompasses a range of topics that will introduce you to the core concepts and techniques prevalent in the field of data mining. These include:

- Association Rules: Understand the principles behind identifying rules that highlight relationships between seemingly independent data in a database.
- Clustering: Learn about grouping a set of objects in such a way that objects in the same group are more similar to each other than to those in other groups.
- Classification: Gain knowledge on the procedures for identifying the predefined class of a new observation.
- Text Mining: Equip yourself with the skills needed to analyze and interpret

large collections of text data to extract meaningful information.

- Data Mining Applications: Explore the various practical applications of data mining across different industries and sectors.

參考書目

Pang-Ning Tan, Michael Steinbach, Vipin Kumar, Introduction to Data Mining, Addison Wesley

課程內容大綱

週次	日期	課程內容
1	1-Sep	Introduction
2	8-Sep	Overview and Data
3	15-Sep	Overview and Data
3	15-Sep	Lab For Data Exploration And Management (Make up for Mid-Autumn Festival)
4	22-Sep	Classification
5	29-Sep	Classification (本週不直播上課，將有課程錄影與學習進度，請學生自行學習。)
6	6-Oct	Mid-Autumn Festival
7	13-Oct	Text Mining & Project Progress Report
8	20-Oct	Lab 2

9	27-Oct	Text Mining
10	3-Nov	Text Mining
11	10-Nov	DM Clustering
12	17-Nov	DM Clustering & Project Progress Report
13	24-Nov	Association
14	1-Dec	Student Paper Presentation (同時段同步報告)
15	8-Dec	Final Exam
16	15-Dec	Final Demo Presentation

成績評量方式

- Two assignments: 20%
- One short presentation: 10%
- One project: 25%
- One exam: 35%
- Class participation (in or after class): 10%

課程要求

- 建議學生需已修過Python程式設計、有基本機率概念
- 本課程期末專題採分組開發，請審慎評估可投入的時間在選課，若需退選最晚須於第十週以前退選，以避免影響同組修課同學之權益。

協同老師負責工作

衛星課程的協同老師，需要獨立完成該盟校所有學生的評分，包含考試卷批改、分組簡報評分、期末專題評分等等，各盟校評分獨立作業。

【資料探勘與應用】的協同老師，不需要同步跟課，但是需要找尋適合的助教以及協助下列評分，也因此盟校至少要給予一半的授課學分數，若該校學生除了主導課程老師的上課內容外，還需要協同老師另外補充上課，則需給予更高的學分數。

此課程需要協助的評分事項有：

- 兩個作業 (占學期成績20%)
- 12/1一週的論文閱讀報告 (占學期成績10%) 由各校協同教師負責，教師可以透過線上報告或是實體到教室報告，報告形式以及報告評分方法由各校協同教師自訂。
- 12/8 期末考試，考卷會提前放在NTU Cool上，由協同教師自行決定是否要在電腦教室考試、或是在教室中紙筆作答。
- 12/15 期末報告，教師可自訂評分方法。

總分：協同教師可以決定該課程是要用百分、等地、或通過 / 非通過的成績。