主導課程四：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

班級人數：1200人 (保留200人給清大，聯盟學校平均每校約45人)

開課級別：研究所課程（開放全校大三(含)以上選課）

授課語言：英語授課　　同步遠距上課時間：星期一 9:00~12:00

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

Week	Date	DM (16-Week Fashion)
1	2-Sep	Introduction
2	9-Sep	Overview and Data
3	16-Sep	Overview and Data
4	23-Sep	Lab 1
5	30-Sep	Classification
6	7-Oct	Classification
7	14-Oct	Text Mining
8	21-Oct	Text Mining
9	28-Oct	Lab 2
10	4-Nov	DM Clustering & Project Progress Report
11	11-Nov	DM Clustering
12	18-Nov	Association & Project Progress Report
13	25-Nov	Association
14	2-Dec	Final Exam (同時段同步考試)
15	9-Dec	Student Paper Presentation (同時段同步報告)
16	16-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%