Data mining provides useful tools for the analysis, understanding and extraction of useful information from huge databases. These techniques are used in business, finance, medicine and engineering. This course will introduce the techniques used in data mining. Topics will include clustering, classification, estimation, forecasting, statistical analysis and visualization tools.

Techniques for Data Mining 數據採集技術

Grade Descriptor:

A

EXCELLENT – exceptionally good performance and far exceeding expectation in all or most of the course learning outcomes; demonstration of superior understanding of the subject matter, the ability to analyze problems and apply extensive knowledge, and skillful use of concepts and materials to derive proper solutions.

B+

B

GOOD – good performance in all course learning outcomes and exceeding expectation in some of them; demonstration of good understanding of the subject matter and the ability to use proper concepts and materials to solve most of the problems encountered.
有關等級說明的資料，請參閱英文版本。

C
FAIR – adequate performance and meeting expectation in all course learning outcomes; demonstration of adequate understanding of the subject matter and the ability to solve simple problems.

有關等級說明的資料，請參閱英文版本。

D
MARGINAL – performance barely meets the expectation in the essential course learning outcomes; demonstration of partial understanding of the subject matter and the ability to solve simple problems.

有關等級說明的資料，請參閱英文版本。

F
FAILURE – performance does not meet the expectation in the essential course learning outcomes; demonstration of serious deficiencies and the need to retake the course.

有關等級說明的資料，請參閱英文版本。

Equivalent Offering:
Units: 3 (Min) / 3 (Max) / 3 (Acad Progress)
Grading Basis: Graded
Repeat for Credit: N
Multiple Enroll: N
Course Attributes: MSc Computer Science
MPhil-PhD Computer Sci & Erg
Topics:

COURSE OUTCOMES

Learning Outcomes:

At the end of the course of studies, students will have acquired the ability to
1. describe the whole data mining process,
2. identify the objectives and tasks of data mining problems,
3. select and apply the appropriate techniques to solve data mining problems.

Course Syllabus:

Data mining provides useful tools for the analysis, understanding, and extraction of useful information from huge databases. These techniques are used in business, finance, medicine, and engineering. This course will introduce the techniques used in data mining. Topics will include clustering, classification, estimation, forecasting, statistical analysis, and visualization tools.

Assessment Type:

- Essay test or exam: 60%
- Others: 40%

Feedback for Evaluation:

1. Course evaluation and questionnaire
2. Question-and-answer sessions during class
3. Student consultation during office hours or online

Required Readings:

To be provided by course teacher.

Recommended Readings:

1. Tan, Steinbach and Kumar, *Introduction to Data Mining*, Addison Wesley, 2006
1. ENGG5103
   Acad Organization=CSEG; Acad Career=RPG

**COMPONENTS**

<table>
<thead>
<tr>
<th>Component</th>
<th>Size</th>
<th>Final Exam</th>
<th>Contact</th>
</tr>
</thead>
<tbody>
<tr>
<td>LEC</td>
<td>30</td>
<td>Y</td>
<td>3</td>
</tr>
<tr>
<td>TUT</td>
<td>30</td>
<td>N</td>
<td>1</td>
</tr>
</tbody>
</table>

**ENROLMENT REQUIREMENTS**

1. ENGG5103
   Enrollment Requirement Group:
   For students in MSc Computer Science or MPhil-PhD programmes under Faculty of Engineering or UG Computer Science or UG Computer Engineering;
   Not for students who have taken CMSC5724 or CSCI5180

CAF

<END OF REPORT>