

**City University of Hong Kong
Course Syllabus**

**offered by School of Data Science
with effect from Semester A 2023/24**

Part I Course Overview

Course Title: Data Mining and Knowledge Discovery

Course Code: SDSC8009

Course Duration: One Semester

Credit Units: 3

Level: R8

Medium of Instruction: English

Medium of Assessment: English

Prerequisites:
(Course Code and Title) Nil

Precursors:
(Course Code and Title) Basic Machine Learning Knowledge, Python Programming

Equivalent Courses:
(Course Code and Title) Nil

Exclusive Courses:
(Course Code and Title) Nil

Part II Course Details

1. Abstract

Data mining studies algorithms and computational paradigms that allow computers to find patterns and regularities in dataset, perform predictions and generally improve the performance through interaction with data. It is currently regarded as the key element of a more general knowledge discovery process that deals with extracting useful knowledge from raw data. This course will offer students advanced algorithms for mining various types of complex data, especially imaging data.

The curriculum will start with the classical data mining methods for tabular and graph data and next move into data-driven imaging data mining with advanced algorithms. We will review different model architectures and learning algorithms such as clustering, classification and graph neural networks. We will go into a few research topics including self-supervised machine learning and various real-world applications.

2. Course Intended Learning Outcomes (CILOs)

No.	CILOs	Weighting* (if applicable)	Discovery-enriched curriculum related learning outcomes (please tick where appropriate)		
			A1	A2	A3
1.	Understand essential data mining algorithms	15%	✓		
2.	Understand basics of graph mining	15%	✓		
3.	Understand how to implement data mining algorithms introduced in this course.	20%	✓		
4.	Apply algorithms taught in this course into emerging real-world problems.	20%	✓	✓	✓
5.	Demonstrate novel knowledge extracted from data of considered real problems through utilizing algorithms taught in this course	30%	✓	✓	✓
		100%			

A1: Attitude

Develop an attitude of discovery/innovation/creativity, as demonstrated by students possessing a strong sense of curiosity, asking questions actively, challenging assumptions or engaging in inquiry together with teachers.

A2: Ability

Develop the ability/skill needed to discover/innovate/create, as demonstrated by students possessing critical thinking skills to assess ideas, acquiring research skills, synthesizing knowledge across disciplines or applying academic knowledge to self-life problems.

A3: Accomplishments

Demonstrate accomplishment of discovery/innovation/creativity through producing /constructing creative works/new artefacts, effective solutions to real-life problems or new processes.

3. Teaching and Learning Activities (TLAs)

TLA	Brief Description	CILO No.					Hours/week (if applicable)
		1	2	3	4	5	
Lecture	- large class activity - questions and discussion	✓	✓	✓	✓	✓	39 hours/sem

4. Assessment Tasks/Activities (ATs)

Assessment Tasks/Activities	CILO No.					Weighting*	Remarks
	1	2	3	4	5		
Continuous Assessment: <u>100</u> %							
<u>Group Project</u> A collaborative research project based on taught concepts.	✓	✓	✓	✓	✓	40%	
<u>Individual Assignment</u> Four assignments for testing the understanding of a sub-set of taught concepts and their implementation.	✓	✓	✓	✓		30%	
<u>Take-home Test</u> An open book and notes examination aiming at assessing the understanding of the overall materials and some open questions for demonstrating the capability of the further exploration.		✓	✓	✓	✓	30%	
Examination: <u>0</u> % (duration: , if applicable)						100%	

5. Assessment Rubrics

Applicable to students admitted in Semester A 2022/23 and thereafter

Assessment Task	Criterion	Excellent (A+, A, A-)	Good (B+, B)	Marginal (B-, C+, C)	Failure (F)
1. Group Project	Application of class materials and teamwork	High	Significant	Basic	Not even reaching marginal levels
2. Individual Assignment	Application of class materials	High	Significant	Basic	Not even reaching marginal levels
3. Take-home Test	Understanding of class materials	High	Significant	Basic	Not even reaching marginal levels

Applicable to students admitted before Semester A 2022/23

Assessment Task	Criterion	Excellent (A+, A, A-)	Good (B+, B, B-)	Fair (C+, C, C-)	Marginal (D)	Failure (F)
1. Group Project	Application of class materials and teamwork	High	Significant	Moderate	Basic	Not even reaching marginal levels
2. Individual Assignment	Application of class materials	High	Significant	Moderate	Basic	Not even reaching marginal levels
3. Take-home Test	Understanding of class materials	High	Significant	Moderate	Basic	Not even reaching marginal levels

Part III Other Information (more details can be provided separately in the teaching plan)

1. Keyword Syllabus

- Data Mining Essentials
- Dimensionality Reduction and Visualization
- Clustering and Classification
- Graph Mining
- Introduction to Neural Networks
- Network Embedding
- Graph Neural Networks
- Selected Data Mining Application: Imaging Data Science
- Data Mining Software

2. Reading List

2.1 Compulsory Readings

1.	Pang-Ning Tan, Michael Steinbach, and Vipin Kumar. Introduction to Data Mining (2 nd Edition), Pearson, 2018.
2.	Lecture notes
3.	Journal articles and conference papers selected by the instructor

2.2 Additional Readings

NIL