The Hong Kong University of Science and Technology (Guangzhou)

UG Course Syllabus

[Course Title] Introduction to Data Mining

[Course Code] AIAA 3111

[No. of Credits] 3 credits

[Any pre-/co-requisites] UFUG 2601 OR UFUG 2602 OR DSAA 1001 (C++ Programming OR Data Structure and Algorithm Design OR Introduction to Data Science and Analytics)

Name: Hao Liu

Email: liuh@hkust-gz.edu.cn

Office Hours: Wednesday 14:00-16:00 at Rm 503, E3

Course Description

This course explains the fundamental principles, uses, and technical details of data mining techniques through lectures and real-world case studies. The emphasis is on understanding the business applications of these techniques. Topics include data preprocessing, classification, clustering, association analysis, and anomaly detection. Upon completion, students will be able to critically apply data mining methodologies to uncover valuable insights from large datasets and solve practical, data-driven problems in various industries.

Intended Learning Outcomes (ILOs)

By the end of this course, students should be able to:

1. Demonstrate thorough knowledge of the literature and a comprehensive understanding of scientific methods and techniques relevant to AI.

2. Demonstrate practical skills in building AI systems.

3. Critically apply theories, methodologies, and knowledge to address fundamental questions in AI.

4. Independently pursue research or innovation of significance in AI applications.

Weekly Schedule

Week (date)	Topic	ILOs
1 (Sep. 3)	Introduction & Data Preprocessing	ILO 1, ILO 2
2 (Sep. 10)	Pattern Mining I	ILO 1, ILO 2, ILO 3
3 (Sep. 17)	Pattern Mining II	ILO 1, ILO 2, ILO 3
4 (Sep. 24)	Classification (Basic)	ILO 1, ILO 2, ILO 3

5 (Oct. 11)	Classification (Advanced)	ILO 1, ILO 2, ILO 3
6 (Oct. 15)	*Midterm exam*	ILO 1, ILO 2, ILO 3
7 (Oct. 22)	Clustering I	ILO 1, ILO 2, ILO 3
8 (Oct 29)	Clustering II	ILO 1, ILO 2, ILO 3
9 (Nov. 5)	Anomaly Detection	ILO 1, ILO 2, ILO 3
10 (Nov. 12)	Advanced Deep Learning Approaches in Data Mining I	ILO 1, ILO 2, ILO 3
11 (Nov. 19)	Advanced Deep Learning Approaches in Data Mining II: LLMs in Data Mining	ILO 1, ILO 2, ILO 3
12 (Nov. 26)	Group project presentation	ILO 1, ILO 2, ILO 3, ILO 4
13 (8/12-19/12)	*Final exam*	ILO 1, ILO 2, ILO 3, ILO 4

Assessment and Grading

This course will be assessed using criterion-referencing and grades will not be assigned using a curve. Detailed rubrics for each assignment are provided below, outlining the criteria used for evaluation.

Assessments:

Assessment Task	Contribution to Overall Course grade (%)	Due date
Course participation	10%	Every class
Mid-Term test	25%	27/10/2024 at 10:00 – 11:30 (preliminary)
Writing Assignment	20%	Four assignments
Group Project	20%	A week before the final exam week
Final examination	25%	Final exam week

* Assessment marks for individual assessed tasks will be released within two weeks of the due date.

Mapping of Course ILOs to Assessment Tasks

Program of study 1: BEng in Artificial Intelligence		
Assessed Task	**Mapped ILOs**	**Explanation**
Course participation	ILO1, ILO2, ILO3.	Based on constructive engagement in class discussions, this task assesses students' understanding of foundational concepts in data mining (ILO 1), the ability to apply basic data analysis methods (ILO 2), and the capability to evaluate introductory data mining applications (ILO 3).
Writing Assignment	ILO1, ILO2, ILO3.	This task assesses students' understanding of foundational

		concepts in data mining (ILO 1), the ability to apply basic data analysis methods (ILO 2), and the capability to evaluate introductory data mining applications (ILO 3).
Group Project	ILO1, ILO2, ILO3, ILO4	The group project evaluates students' ability to integrate and apply concepts from data mining (ILO 1, ILO 2, ILO 3) to real-world scenarios. It also emphasizes teamwork, collaboration, and communication skills as students work together to develop and present a comprehensive solution, demonstrating synthesis, critical evaluation, and effective team dynamics (ILO 4).
Mid-term Exam	ILO1, ILO2, ILO3	The mid-term exam evaluates students' comprehensive understanding of the course content covered so far, including foundational principles, analytical methods etc. (ILO 1, ILO 2, ILO 4).
Final-term Exam	ILO1, ILO2, ILO3, ILO4	The final exam assesses the full range of learning outcomes, including students' knowledge of data mining principles and their ability to apply, evaluate, and synthesize these concepts in different scenarios (ILO 1, ILO 2, ILO 3, ILO 4).

Grading Rubrics

1. Course participation (10%):

    - Class Attendance: Attendance in class is mandatory. Arriving more than 10 minutes late will result in a deduction of attendance marks. If you anticipate being late or need to leave early, please inform the instructor in advance. Only sick leave with a valid medical certificate will be accepted as a formal reason for absence.

2. Writing Assignment (20%):

    - Assignments: Assignments must be submitted before the deadline. These tasks are crucial for reinforcing the material covered in lectures.

2. Group Project & Presentation (20%):

- Group Formation: Groups must consist of a maximum of 4 members. You are encouraged to start forming your project/study teams at the beginning of the course.
- Presentation: Each group will deliver a 10-minute presentation followed by a 2-minute Q&A session in class, scheduled for one week before the final exam.
- Project Focus: The project should center on the application of Data Mining to real-world problems.
- Project Topic Selection:

    - Option 1: Students are highly encouraged to form a team and brainstorm issues or problems of personal interest that are also relevant to real-world contexts. This option allows for more creativity and exploration.
    - Option 2: If needed, we will provide a list of potential project topics to help guide your exploration. This list will be released as soon as possible, but no later than the midterm exam.
- Grading Rubrics: The grading rubrics for the group project and presentation will be released after the midterm exam.

3. Exams (50% in total):

- All exams for this course are closed book. The exams will consist of multiple-choice questions, single-answer questions, and other formats designed to test your understanding of the course material.

Final Grade Descriptors:

Grades	Short Description	Elaboration on subject grading description
A	Excellent Performance	Demonstrates a comprehensive grasp of subject matter, expertise in problem-solving, and significant creativity in thinking. Exhibits a high capacity for scholarship and collaboration, going beyond core requirements to achieve learning goals.
B	Good Performance	Shows good knowledge and understanding of the main subject matter, competence in problem-solving, and the ability to analyze and evaluate issues. Displays high motivation to learn and the ability to work effectively with others.
C	Satisfactory Performance	Possesses adequate knowledge of core subject matter, competence in dealing with familiar problems, and some capacity for analysis and critical thinking. Shows persistence and effort to achieve broadly defined learning goals.
D	Marginal Pass	Has threshold knowledge of core subject matter, potential to achieve key professional skills, and the ability to make basic judgments. Benefits from the course and has the potential to develop in the discipline.
F	Fail	Demonstrates insufficient understanding of the subject matter and lacks the necessary problem-solving skills. Shows limited ability to think critically or analytically and exhibits minimal effort towards achieving learning goals. Does not meet the threshold requirements for professional practice or development in the discipline.

Course AI Policy

Allowed and encouraged with proper citations and prompt list.

Communication and Feedback

Assessment marks for individual assessed tasks will be communicated via Canvas within two weeks of submission. Feedback on assignments will include [specific details, e.g., strengths, areas for improvement]. Students who have further questions about the feedback including marks should consult the instructor within five working days after the feedback is received.

Resubmission Policy

Resubmissions are not permitted after the deadline. However, if you wish to appeal your grade or feedback, please refer to the "Communication and Feedback" section for the appropriate procedure.

Required Texts and Materials

Not Required. Recommended reading materials will be introduced during classes.

Academic Integrity

Students are expected to adhere to the university's academic integrity policy. Students are expected to uphold HKUST(GZ)'s Academic Honor Code and to maintain the highest standards of academic integrity. The University has zero tolerance of academic misconduct. Please refer to Regulations for Academic Integrity and Student Conduct for the University's definition of plagiarism and ways to avoid cheating and plagiarism.