[Course Title] Machine Learning with Structured Data

[Course Code] COMP4222

[No. of Credits] 3 credits

[Any pre-/co-requisites]

Name: Yangqiu Song

Email: yqsong@ust.hk

Course Description

This course provides an introduction to learning algorithms for structured data such as sequences, trees, relational databases (such as knowledge bases), and graphs (including graph databases such as large heterogeneous information networks and knowledge graphs, and using programming tools such as Python to implement them for real problems. It will use some of the following practical problems such as graph classification, statistical relational learning, structure prediction models, QA system, etc. as illustrations to demonstrate the power of the learning algorithms.

Intended Learning Outcomes (ILOs)

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

• Explain the basic principles behind machine learning algorithms for structured data.

• Implement programs for structured prediction tasks.

• Formulate machine learning solutions to domain problems.

• Demonstrate the ability to understand the complexity of real-world problems.

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 (%)
Reading notes and assignments	20-30%
Project code and report	20-30%
Final project presentation	10%

Final exam	40%	
Total	100%	

Assessments:

Assessment Task	Contribution to Overall Course grade (%)	Due date
Reading notes and assignments	20-30%	20/12/2024
Project code and report	20-30%	20/12/2024
Final project presentation	10%	20/12/2024
Final exam	40%	20/12/2024

Mapping of Course ILOs to Assessment Tasks

Assessed Task	Mapped ILOs	Explanation
Reading notes and assignments	ILO1, ILO4	This task assesses students' ability to comprehend and explain foundational machine learning principles (ILO1) and apply them to analyze complex real-world scenarios (ILO4).
Project code and report	ILO2, ILO3, ILO4	This task assesses students' ability to develop structured prediction programs (ILO2), formulate domain-specific machine learning solutions (ILO3), and evaluate real-world problem complexity through implementation (ILO4).
Final project presentation	ILO4	This task assesses students' ability to articulate the complexity of real-world problems (ILO4) and present viable machine learning strategies in a clear, professional manner.
Final exam	ILO1, ILO2, ILO3, ILO4	This task assesses students' ability to demonstrate theoretical knowledge of machine learning algorithms (ILO1), solve structured prediction tasks (ILO2), design problem-specific solutions (ILO3), and critically assess real-world application complexities (ILO4).

Grading Rubrics

The reading notes are graded based on explaining the paper's relevance to course content, highlighting two strengths and one weakness in the review, and posing critical questions for the authors.

The project code and report are graded based on clarity of problem introduction, methodological justification, experimental results, and discussions incorporating case studies, error analysis, and unique insights into the research problem.

The presentation grading rubric evaluates **content clarity, method description, experiments/components, insight sharing, and time management** on a 1-5 scale, with final scores averaged across instructor and TAs.

The grading criteria for assignments and final exams will be evaluated on a case-by-case basis.

Final Grade Descriptors:

Grades	Short Description	Elaboration on subject grading description
A	Excellent Performance	Students master structured data algorithms, implement complex tasks precisely, design innovative solutions, and critically analyze real-world complexities, exceeding expectations in both theory and practice.
B	Good Performance	Students grasp core principles, implement tasks reliably, propose logical solutions, and articulate real-world challenges effectively, with minor gaps in technical depth or innovation.
C	Satisfactory Performance	Students understand basic concepts, complete simple tasks with occasional errors, develop generic solutions, and partially address complexities, meeting minimum requirements.
D	Marginal Pass	Students show limited comprehension, struggle with functional implementation, propose incomplete strategies, and superficially engage with real-world problems, falling short of expectations.
F	Fail	Students fail to explain principles, produce non-functional code, lack coherent solutions, and disregard practical complexities, demonstrating insufficient engagement with course objectives.

Course AI Policy

Students may use generative AI tools to assist with project design, coding, or analysis, provided all AI-generated content is properly cited and its contributions are transparently explained in final submissions.

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 scoring criteria breakdowns, identified strengths, areas needing improvement, and detailed point deductions for each evaluated component. 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

Students may submit assignments multiple times prior to the deadline, with the final submission automatically designated for grading. Late submission will not be accepted except in exceptional circumstances requiring prior approval from the instructor.

Required Texts and Materials

N/A

Academic Integrity

Students are expected to adhere to the university's academic integrity policy. Students are expected to uphold HKUST's Academic Honor Code and to maintain the highest standards of academic integrity. The University has zero tolerance of academic misconduct. Please refer to <u>Academic Integrity | HKUST – Academic Registry</u> for the University's definition of plagiarism and ways to avoid cheating and plagiarism.

Additional Resources

N/A