The Hong Kong University of Science and Technology

UG Course Syllabus

Capstone Project for Data Science

MATH4995

3 credits

Prerequisites: MATH 3322 AND MATH 3332

Name: Jianfeng Cai

Email: jfcai@ust.hk

Office Hours: MoFr 9:00-10:00 or by appointment or walk-in.

Course Description

This project-based course bridges the gap between mathematical theory and cutting-edge artificial intelligence, designed for students in data science and technology. You will analyze seminal papers from top AI conferences, applying your knowledge of advanced optimization, numerical linear algebra, and probabilistic modeling to deconstruct the foundations of key architectures like Transformers, Graph Neural Networks, and Generative AI models. The goal is to build a first-principles understanding of why these systems work. The course culminates in a final project where you will deeply engage with a research topic. Projects may involve implementing and extending published algorithms, conducting a rigorous theoretical analysis of their properties, or surveying a subfield to synthesize new insights. This approach empowers you to critique, improve, and invent AI methods from a foundational perspective. Strong programming skills in Python (with PyTorch/TensorFlow) are highly valuable for the experimental aspects of the course.

Intended Learning Outcomes (ILOs)

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

1. Explain the mathematical foundations (such as optimization and matrix computation) behind modern AI models.

2. Analyze research papers from top AI conferences to identify their key algorithms and contributions.

3. Implement AI algorithms using Python frameworks (e.g., PyTorch or TensorFlow) to reproduce experimental results.

4. Conduct a project that surveys a specific topic or applies mathematical theory to solve a data science problem.

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
Meeting attendance	20%	
Midterm oral presentation	20%	17/10/2025
Final oral presentation	30%	28/11/2025
Final written report	30%	24/12/2025

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

Mapping of Course ILOs to Assessment Tasks

Assessed Task	Mapped ILOs	Explanation
Meeting Attendance	ILO1, ILO2, ILO3, ILO4.	This task assesses students' continuous engagement with the course material. Through regular discussions, students must articulate their understanding of mathematical concepts (ILO1), report on their reading of current literature (ILO2), provide updates on their coding progress (ILO3), and demonstrate consistent management of their project milestones (ILO4).
Midterm oral presentation	ILO1, ILO2, ILO4.	This presentation assesses the students' ability to define the scope of their project (ILO4) by synthesizing background research. Students are expected to explain the relevant mathematical theories (ILO1) and critically analyze the key algorithms found in the research papers (ILO2) that will form the basis of their proposed work.
Final oral presentation	ILO1, ILO3, ILO4.	This task assesses the culmination of the project (ILO4). Students must demonstrate the successful implementation of their AI algorithms (ILO3), present their experimental results, and use mathematical foundations (ILO1) to interpret and justify why the model performed as it did.
Final written report	ILO1, ILO2, ILO3, ILO4.	The final report is a comprehensive assessment requiring students to synthesize all learning outcomes. It requires a detailed written explanation of the mathematical optimization involved (ILO1), a formal literature review of related works (ILO2), documentation of the Python implementation and reproducibility steps (ILO3), and a structured

		discussion of the project's methodology and conclusions (ILO4).

Grading Rubrics

1. Mid-term Oral Report Rubric: Problem Definition & Context (25%), Methodology & Planning (30%), Preliminary Progress (25%), Presentation Skills (20%).
2. Final Oral Report Rubric: Content & Results (35%), Analysis & Conclusion (30%), Q&A Response (20%), Organization & Flow (20%).
3. Final Written Report: Executive Summary / Abstract (10%), Technical Content & Depth (30%), Critical Thinking & Discussion (25%), Structure & Organization (15%), Writing Style & Mechanics (20%).

Final Grade Descriptors:

Letter grades will be assigned according to the final calculated score.

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

To ensure academic integrity, the use of Generative AI tools (e.g., ChatGPT, Grammarly) is restricted to editorial assistance only. You may use these tools to correct grammar, improve sentence structure, and refine the clarity of your English.

However, you are strictly prohibited from using AI to generate the substance of your work. AI must not be used to write entire sections, conduct analysis, or formulate your conclusions. The intellectual content, reasoning, and data analysis must be entirely your own.

Communication and Feedback

Students are expected to meet with the instructor or TAs once a week. Assessment marks for individual assessed tasks will be communicated via Canvas within two weeks of submission. 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

Resubmission is not allowed.

Required Texts and Materials

None.

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 Academic Integrity | HKUST – Academic Registry for the University's definition of plagiarism and ways to avoid cheating and plagiarism.