THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY (GUANGZHOU)
Data Science and Analytics Thrust

DSAA 4011

Advanced Machine Learning and Deep Learning

Term	Fall 2026–27	**Credits**	Four
Instructor	Yuwen Huang	**Meeting dates**	September 1–December 7, 2026
Email	yuwenhuang@hkust-gz.edu.cn	**Office hours**	Tuesday 11:00 a.m.–12:00 noon, W4 L4 422
Prerequisite	DSAA 2011 or DSAA 2012		
Lecture	Thursday, 9:00–11:50 a.m., W1-233	**LA01**	Friday, 3:00–3:50 p.m., E1-150

Course description

This student guide supplements the official course record. It explains the Fall 2026–27 teaching plan, assessment, and course policies.

DSAA 4011 studies how local relationships and observations can be combined to answer questions about a whole system. Students first express a probability model as a factor graph and use message passing to compute exact probabilities on trees. When exact calculation is too costly, students study information measures, energy-based models, sampling, learning, and approximate inference. The course then examines neural belief propagation, approximate message passing, multiscale networks, diffusion models, attention, and language models. Laboratories use small, checkable examples before larger experiments. Students test boundary cases and support conclusions with reproducible evidence.

Learning goals for this offering

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

1. Construct factor graphs and derive exact message-passing algorithms on trees.
2. Use information measures and energy-based representations to analyze probabilistic models.
3. Implement and diagnose sampling, parameter learning, structure learning, and approximate inference.
4. Compare fixed, approximate, and learned inference procedures under controlled assumptions.
5. Explain the relationship between structured inference and selected modern neural or generative models without conflating their semantics.
6. Design, test, and defend reproducible experiments under numerical, statistical, and computational constraints.

Scope and foundations

The course introduces the probability, linear algebra, optimization, and numerical ideas needed for each topic. It is designed to be self-contained for students who have completed DSAA 2011 or DSAA 2012. In probabilistic message passing, each message has a defined probabilistic meaning and, on a tree, produces an exact result. A neural network may also pass vectors between nodes, but those vectors need not represent probabilities. Therefore, the exactness result for probabilistic message passing does not automatically apply to a neural network.

Teaching team

Role	Name	Responsibilities
Instructor	Yuwen Huang	Lectures, coordination, assessment design, examinations, and final academic decisions.
Teaching Assistant	Yuhan Chen	LA01, laboratory support, and midterm marking. Office hour: Friday, 4:00–5:00 p.m., E1, 5F, workstation 333.

Weekly schedule

Week	Lecture topic
1	History and factor-graph foundations
2	Exact inference on trees
3	Information, energy, and probabilistic models
4	Markov chain Monte Carlo
5	Learning with hidden variables
6	Approximate inference
7	Learning graph structure from data
8	Midterm examination and guided error review
9	Neural belief propagation
10	Sparse recovery with approximate message passing
11	Multiscale networks and U-Net
12	Diffusion models
13	Attention
14	Language models and course synthesis

Laboratories and the individual inquiry project

Weekly laboratories are guided practice. They do not form a separate project and do not directly contribute to the course grade. Students use small examples to check a calculation, implement a method, and examine a failure case. Coursework briefs define all assessed submissions.

The Individual Inquiry Project consists of Coursework 3 and Coursework 4; it is not an additional laboratory project. Students complete the project individually outside the scheduled laboratory. Part I states a focused question, implements a reference method, and designs a fair comparison. Part II extends the same study, evaluates several justified boundary cases or shifted datasets, and presents a concise report and brief defense. One result on one dataset is not sufficient. Students must explain what changes, when a method fails, and which evidence supports the conclusion.

Students select one approved track: fixed versus learned belief propagation; approximate message passing versus a learned iterative method; a local denoiser versus U-Net or a diffusion denoiser; or an instructor-approved question.

Assessment

Assessment	Weight	Timing
Coursework 1: Model audit	10%	Weeks 1–3
Coursework 2: Inference under stress	10%	Weeks 4–7
Individual Inquiry Project, Part I	10%	Weeks 9–11
Individual Inquiry Project, Part II	10%	Weeks 12–14
Midterm examination	20%	Week 8
Final examination	40%	Examination period
Total	**100%**	

Common coursework rubric

Criterion	Weight
Technical correctness	35%
Reasoning and diagnosis	30%
Implementation and evidence	25%
Clarity and reproducibility	10%
Total	**100%**

Each coursework task contains 90 marks of required work and a 10-mark optional extension, for a maximum of 100. Originality means making and justifying an independent choice. Innovation means extending the supplied baseline through a new test, comparison, modification, or failure analysis. Students do not need to invent a new algorithm.

Examinations and course policies

Examinations

Both examinations are closed book. The midterm takes place in Week 8 and covers Weeks 1–7. The Week 8 laboratory uses a new example to show how to identify whether an error comes from the model, the question being answered, the algorithm, the numerical calculation, or the evidence used to justify a conclusion. It does not reproduce a midterm question. The final examination covers Weeks 1–14 and takes place during the University examination period. Canvas will confirm the duration, location, and permitted materials.

Preparation, class, and laboratory

Before class, review the assigned notes and complete the stated readiness questions. Class sessions combine concise explanations, worked examples, short prediction questions, and discussion; the balance varies by topic. Laboratories provide guided practice with small, checkable examples.

Students need basic Python. Starter materials introduce Jupyter, NumPy, SciPy, NetworkX, and PyTorch when needed. Core tasks use small models, generated data, or controlled simulations; they do not require costly computing hardware.

Language models and coding agents

Coursework and laboratories permit language models and coding agents; examinations prohibit them. You remain responsible for every model, derivation, test, program, result, and claim. If you use AI in graded work, submit the complete relevant thread, tool and model information, your edits and decisions, and the calculations, tests, or experiments used for verification. Do not submit unrelated private or confidential material. If you do not use AI, state this in the submission. AI-detection scores do not determine misconduct.

Independent work and integrity

Students may discuss concepts, documentation, and general strategies. Each student must submit independent derivations, code, experiments, and explanations. Acknowledge substantial help and cite borrowed data, code, text, and ideas. HKUST(GZ)'s Academic Honor Code and the Regulations for Academic Integrity and Student Conduct apply.

Deadlines, feedback, and grade review

Announcements, materials, and submissions use Canvas. Use email or a private Canvas channel for personal matters. Results and feedback normally appear within two weeks when institutional schedules permit. An extension requires a documented, well-justified circumstance and written instructor approval. Without approval, apply the following multiplier to the raw mark for each started 24-hour period: on time, 1.0; up to 24 hours late, 0.8; 24–48 hours, 0.6; 48–72 hours, 0.4; 72–96 hours, 0.2; more than 96 hours, 0.

Submit a reasoned grade-review request through Canvas within five working days after feedback is released. Name one published rubric area and identify a specific marking error. The reviewed mark may rise, fall, or stay unchanged. Resubmission after marking is not normally permitted; documented exceptions follow University policy.

For approved learning or assessment arrangements, contact the University's student support service and the instructor early. Yuhan Chen leads LA01. The teaching team may assist with invigilation; the instructor approves all marking guides and final grading decisions.

Core references

Factor graphs	Kschischang, Frey, and Loeliger (2001), Factor Graphs and the Sum-Product Algorithm
Information and inference	MacKay (2003), Information Theory, Inference, and Learning Algorithms
Graphical models	Koller and Friedman (2009), Probabilistic Graphical Models
Variational inference	Wainwright and Jordan (2008), Graphical Models, Exponential Families, and Variational Inference
Structure learning	Chow and Liu (1968), Approximating Discrete Probability Distributions with Dependence Trees
Approximate messages	Donoho, Maleki, and Montanari (2009), Message-Passing Algorithms for Compressed Sensing
Multiscale networks	Ronneberger, Fischer, and Brox (2015), U-Net: Convolutional Networks for Biomedical Image Segmentation
Diffusion	Ho, Jain, and Abbeel (2020), Denoising Diffusion Probabilistic Models
Attention	Vaswani et al. (2017), Attention Is All You Need