THE CHINESE UNIVERSITY OF HONG KONG
Print Course Catalog Details

Academic Org: Div of Computer Science & Engg – Subject: AI: Systems & Tech

Course: AIST5050 **Course ID:** 015392 **Eff Date:** 2026-07-01 **Crse Status: Active** **Apprv. Status:** Approved **[**New Course**]**
Optimization for Large Language Models 大語言模型優化

This course introduces the advanced optimization principles that power modern large language model (LLM) systems across their full lifecycle. This course begins with training time adaptive optimization methods, including Adaptive Moment Estimation (Adam), Adaptive Moment Estimation with decoupled weight decay (AdamW), Adafactor, Lion, and Muon, along with learning rate schedules and gradient clipping. This course also covers parameter efficient fine tuning, specifically Low Rank Adaptation (LoRA) and Quantized Low Rank Adaptation (QLoRA), for settings with tight memory and compute budgets.

Next, post training with feedback will be introduced, including Reinforcement Learning from Human Feedback (RLHF) using Proximal Policy Optimization with a Kullback Leibler divergence constraint (PPO), Direct Preference Optimization (DPO), and emerging variants such as Group Relative Policy Optimization (GRPO). Throughout, the course emphasize objective design, stability, and compute efficiency.

Finally, the course frame inference time optimization as a budgeted decision problem, covering best of N selection and self consistency, Tree of Thoughts, contrastive decoding, speculative decoding, and lookahead decoding. this course also introduces in context learning and its interpretation as implicit Bayesian inference and gradient descent in context. To connect modalities, the course includes a short diffusion unit to contrast optimization and sampling strategies.

The course is hands on and research oriented. Students will learn the theoretical principles of these techniques, run small scale ablations, match floating point operations when comparing methods, and present a final project.

Advisory: Students are expected to have taken CSCI3320 or ENGG 5501 or equivalent.

本科介紹支撐現代大型語言模型（LLM）全生命週期的高級優化原理。我們首先從訓練階段的自適應優化方法入手，涵蓋自適應動量估計（Adam）、帶解耦權重衰減的自適應動量估計（AdamW）、Adafactor、Lion 與 Muon，並配合學習率排程與梯度裁剪。本科亦將討論參數高效微調，重點包括低秩適配（LoRA）與量化低秩適配（QLoRA），以應對記憶體與算力資源受限的情境。

隨後將介紹帶回饋的後訓練技術，包括基於人類回饋的強化學習（RLHF），其實現採用帶 Kullback–Leibler 散度約束的近端策略優化（PPO），以及直接偏好優化（DPO）與新近方法（如組相對策略優化，GRPO）。全過程強調目標函數設計、穩定性與算力/計算效率。

最後，介紹將推理階段優化建模為一個受預算約束的決策問題，討論 Best-of-N 選擇與自一致性（self-consistency）、思維樹（Tree of Thoughts）、對比解碼（contrastive decoding）、投機解碼（speculative decoding）與前瞻解碼（lookahead decoding）。同時介紹上下文學習（In-Context Learning, ICL），並將其解釋為一種隱式貝葉斯推斷與「在上下文中的梯度下降」。為銜接多模態，我們另設一個簡短的擴散模型單元，用以對比優化與取樣策略。

本科重視實作與研究導向。學生將掌握上述技術的理論原理，開展小規模消融實驗，在方法比較時匹配浮點運算量（FLOPs）以確保公平性，並完成期末專題之展示。

建議：學生應曾修讀CSCI3320或ENGG5501或同等科目

Grade Descriptor: A

EXCELLENT – exceptionally good performance and far exceeding expectation in all or most of the course learning outcomes; demonstration of superior understanding of the subject matter, the ability to analyze problems and apply extensive knowledge, and skillful use of concepts and materials to derive proper solutions.

有關等級說明的資料，請參閱英文版本。

B

GOOD – good performance in all course learning outcomes and exceeding expectation in some of them; demonstration of good understanding of the subject matter and the ability to use proper concepts and materials to solve most of the problems encountered.

有關等級說明的資料，請參閱英文版本。

C

FAIR – adequate performance and meeting expectation in all course learning outcomes; demonstration of adequate understanding of the subject matter and the ability to solve simple problems.

有關等級說明的資料，請參閱英文版本。

D

MARGINAL – performance barely meets the expectation in the essential course learning outcomes; demonstration of partial understanding of the subject matter and the ability to solve simple problems.

有關等級說明的資料，請參閱英文版本。

F

FAILURE – performance does not meet the expectation in the essential course learning outcomes; demonstration of serious deficiencies and the need to retake the course.

有關等級說明的資料，請參閱英文版本。

Equivalent Offering:
Units: 3 (Min) / 3 (Max) / 3 (Acad Progress)
Grading Basis: Graded
Repeat for Credit: N
Multiple Enroll: N
Course Attributes: MSc Computer Science
MPhil-PhD Computer Sci & Erg

Topics:

COURSE OUTCOMES

Learning Outcomes:

At the end of the course of studies, students will be able to:
1. Implement at least four optimizers (AdamW, Adafactor, Lion, Muon) with LR schedules and clipping and justify choices via ablations.
2. Fine-tune a model with LoRA/QLoRA and report compute-normalized quality and memory.
3. Reproduce a small-scale RLHF baseline and a DPO run; evaluate with preference win-rates and calibration.
4. Prototype GRPO-style training on a reasoning toy task; analyze sample cost vs. gains.
5. Build a decode-time optimizer that meets an accuracy/latency target using best-of-N, ToT, or contrastive decoding; quantify improvements.

Course Syllabus:

Week 1: Optimization across the LLM lifecycle. Taxonomy of objectives (training vs. post-training vs. decoding). An Overview of Optimization.
Week 2: Optimizers I — Adam/AdamW, Adafactor Decoupled weight decay vs. L2; second-moment factorizations; gradient clipping.
Week 3: Optimizers II — Lion, Muon; LR warmup/cosine; gradient clipping Sign-momentum (Lion); orthogonality-flavored Muon; warm restarts; exploding-gradients & clip-norm.
Week 4: Optimizers III: Theory of Adaptive Methods: convergence analysis of adaptive methods under different conditions (convex, PL, smooth non-convex), including iteration complexity and sample complexity.
Week 5: Efficient fine-tuning (PEFT): LoRA and QLoRA Rank choice, adapter placement.
Week 6: Post-training I — Reinforcement learning from human feedback (RLHF) (KL-regularized PPO) InstructGPT pipeline; reward-model fit; KL trust-region view.

Week 7: Post-training II — DPO and variants: Preference fitting as logistic risk; reference-model effects; PPO-style RLHF.
Week 8: Post-training III — GRPO & resource-aware RL: Group-relative baselines, critic-free PPO; sampling costs and scaling; current analyses & tweaks (CPPO).
Week 9: Decoding I — sampling controls and contrastive methods Temperature, top-k; contrastive search vs. contrastive decoding.
Week 10: Decoding II — test-time compute: Self-consistency (best-of-N), Tree-of-Thoughts search
Week 11: Serving-time acceleration: Speculative decoding (draft-then-verify) and lookahead decoding (parallel exact steps)
Week 12: In-Context Learning (ICL) — statistical and optimization views ICL as implicit Bayes; transformers emulating GD; synthetic setups that reproduce ICL phenomena.
Week 13: Diffusion models (optimization and sampling) denoising diffusion probabilistic model (DDPM) objective; DDIM acceleration; DPM-Solver

Assessment Type:

Homework or assignment	: 60%
Project	: 40%

Feedback for Evaluation:

1. Course evaluation and questionnaire
2. Question-and-answer sessions during class
3. Student consultation during office hours or online

Required Readings:

To be provided by course teacher.

Recommended Readings:

1. Loshchilov & Hutter (2017). Decoupled Weight Decay Regularization (AdamW). ICLR.
2. Chen et al. (2023). Symbolic Discovery of Optimization Algorithms (Lion). arXiv.
3. Jordan (2024). Muon: An optimizer for hidden layers in neural networks. Blog.
4. oshchilov & Hutter (2017). SGDR: Stochastic Gradient Descent with Warm Restarts. ICLR.
5. Hu et al. (2022). LoRA: Low-Rank Adaptation of Large Language Models. ICLR.
6. Dettmers et al. (2023). QLoRA: Efficient Finetuning of Quantized LLMs. NeurIPS.
7. Rafailov et al. (2023). Direct Preference Optimization (DPO). NeurIPS/OpenReview.
8. Shao et al. (2024). DeepSeekMath (introduces GRPO). arXiv.
9. Wang et al. (2023). Self-Consistency Improves Chain-of-Thought Reasoning. ICLR.

10. Xie et al. (2022). An Explanation of In-Context Learning as Implicit Bayesian Inference. ICLR.
11. Ho, Jain, Abbeel (2020). DDPM. NeurIPS.

OFFERINGS

1. AIST5050 — Acad Organization=CSEGV; Acad Career=RPG

COMPONENTS

LEC : Size=60; Final Exam=N; Contact=3
TUT : Size=60; Final Exam=N; Contact=1

ENROLMENT REQUIREMENTS

1. AIST5050

Enrollment Requirement Group:

For students in MSc Computer Science; or
For students in MPhil-PhD Computer Science & Engineering; or
For students in UG AISTN or CDASN or CENGN or CSCIN

New Enrollment Requirement(s):

Other Requirement = For students in MSc Computer Science; or
For students in MPhil-PhD Computer Science & Engineering; or
For students in UG AISTN or CDASN or CENGN or CSCIN

Additional Information

VTL-Onsite face-to-face hrs 0
VTL-Online synch. hrs 0
VTL-Online asynch. hrs 0

< E N D O F R E P O R T >