

INF 385T: Post-Training of Large Language Models

Fall 2026 / School of Information / The University of Texas at Austin

Instructor	Chan Young Park
Email	park@utexas.edu
Office	UTA 5.442
Office hours	Fridays 10–11am. Please email me in advance to confirm.
Teaching Assistant	Arjun Nagendr (aa229243@my.utexas.edu) Office hours: By appointment; Zoom/UTA. Scheduling information will be posted on Canvas.
Class meetings	Wednesdays 12–3pm
Classroom	UTA 1.212
Modality	In person. Attendance in the classroom is expected. Sessions are not livestreamed.
Canvas	https://utexas.instructure.com/courses/1454896
Unique number	30845

Course Description

A pretrained language model is not yet a useful assistant. Much of what makes modern LLMs useful happens after pretraining, including supervised fine-tuning, parameter-efficient adaptation, reward modeling, reinforcement learning from human and AI feedback, direct preference optimization, and reinforcement learning with verifiable rewards.

This is a hands-on course. Over the course of the semester, you will fine-tune open-weight models on GPUs, build evaluation suites for your own tasks, construct preference data, and train models with DPO and GRPO. We will use lecture time to develop the methods and the research behind them and the labs are where we run the models, inspect the results, and see what breaks.

Prerequisites

Required: Graduate standing; proficiency in Python; familiarity with basic machine learning concepts; and comfort with the command line. Prior coursework or equivalent experience in machine learning or NLP is expected.

Helpful but not required: Prior experience with PyTorch, the Hugging Face ecosystem, or HPC clusters. We will introduce the tools needed for the course, but students should be comfortable learning new ML libraries and working in a computational environment.

How this course works

Most Wednesdays will be divided between lecture and a hands-on lab or workshop. The lecture will introduce the methods, concepts, and research behind that week's topic. The lab or workshop will provide time either to apply those ideas directly or to connect them to the group project through experimentation and discussion. A major component of the course is a group project. You will work in teams of three to post-train a model for a task of your choosing, making decisions about data, evaluation, training methods, and iteration along the way.

Learning Outcomes

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

Learning outcome	Assessed by
Decide whether a given task calls for prompting, retrieval, or post-training, and justify the decision with evidence	Project proposal, concept check quiz, final report

Learning outcome	Assessed by
Design and implement a task-specific evaluation suite, including an LLM-as-a-judge component, and assess its validity.	Evaluation labs, project milestones, final report
Implement supervised fine-tuning and parameter-efficient fine-tuning (LoRA/QLoRA) on open-weight models	HW1, project
Construct preference data and train a model with DPO; explain the relationship between RLHF and DPO	HW2, concept check quiz
Understand reward modeling, PPO, RLAIIF, and GRPO at a level sufficient to read current papers and choose among methods	Concept check quiz, in-class lab
Diagnose training runs and common failure modes	HW1, HW2, final report
Communicate methods, results, and limitations to a technical audience	Final presentation, final report

Course Materials

There is no required textbook. Readings will consist of research papers, documentation, and lab notebooks made available through Canvas at no cost. Papers discussed in the course may include:

- DeepSeek-AI, “DeepSeek-R1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning,” 2025.
- Rafailov et al., “Direct Preference Optimization,” NeurIPS 2023.
- Ouyang et al., “Training language models to follow instructions with human feedback,” NeurIPS 2022.
- Hu et al., “LoRA: Low-Rank Adaptation of Large Language Models,” ICLR 2022.
- Bai et al., “Constitutional AI: Harmlessness from AI Feedback,” Anthropic, 2022.

Required Technology

- A laptop capable of SSH access and running a modern browser. Model training will take place on remote GPUs rather than on your laptop.
- A TACC (Texas Advanced Computing Center) account, which will be set up at the beginning of the course. Compute for coursework will be provided through a class allocation at no cost to you.
- Access to course-provided remote model services, with onboarding instructions provided by the instructor, for activities such as model comparison, LLM-as-a-judge evaluation, and synthetic data generation.

Getting help with technology: If you run into problems with course infrastructure, including TACC, SLURM, or the course-provided remote model services, please ask for help through Canvas or during TA office hours. You are encouraged to ask questions whenever technical setup or infrastructure is getting in the way of making progress.

Course Schedule

Subject to change. Any changes will be announced in class and posted to Canvas. Written work is due at 11:59 pm on the listed date; in-class assessments and presentations take place during class.

Wk	Date	Lecture	Lab / Workshop	Due
1	Aug 26	Intro to post-training.	TACC account setup; base vs. instruction-tuned model comparison	HW0 out
2	Sep 2	LLM landscape. Prompting and in-context learning as baselines	Chat templates, in-context learning, and cross-model comparison; project team formation	
3	Sep 9	Evaluation I: task metrics, held-out design, what a benchmark can and cannot tell you	Proposal workshop	HW0 due
4	Sep 16	Evaluation II: LLM-as-a-judge, rubric design, judge bias and agreement	Build and validate a judge pipeline	

Wk	Date	Lecture	Lab / Workshop	Due
5	Sep 23	Supervised fine-tuning: objective, chat templates, hyperparameters	First SFT run on TACC	Proposal due HW1 out
6	Sep 30	Parameter-efficient fine-tuning: LoRA, QLoRA	LoRA rank ablation and cost comparison	
7	Oct 7	Data for post-training: curation, synthetic generation, distillation	Synthetic data generation and filtering	HW1 due
8	Oct 14	RLHF foundations: Bradley-Terry reward models, PPO	Train and probe a reward model	
9	Oct 21	Direct preference optimization (DPO)	Preference labeling and guided DPO workflow	HW2 out
10	Oct 28	RLAIF and Constitutional AI. Scalable oversight and its limits	AI-feedback pipeline; project checkpoint	
11	Nov 4	Verifiable rewards and GRPO. Reasoning-oriented post-training	Guided GRPO lab	HW2 due
12	Nov 11	Failure modes: reward hacking, sycophancy, catastrophic forgetting, safety regressions	Concept check quiz (in class, handwritten, closed-device), then project work	
13	Nov 18	Project presentations	Presentation feedback and discussion	Final presentation
—	Nov 25	No class (Thanksgiving break)		
14	Dec 2	Feedback-based project revision and refinement		Final report due Dec 7

Assignments and Major Requirements

The opening labs establish inference and prompting baselines. The individual homework then progresses from reproducible compute practice to supervised fine-tuning and preference optimization. The project asks your team to integrate those ideas on a task you choose.

HW0. TACC inference and resource accounting (released Aug 26, due Sep 9, 5%)

Use a provided SLURM workflow to reproduce course-designated model inference on TACC, inspect job outputs and accounting information, and estimate compute use. The emphasis is on reproducibility and responsible use of shared resources.

HW1. Supervised fine-tuning and LoRA (released Sep 23, due Oct 7, 15%)

Fine-tune a course-designated open-weight model with LoRA on a provided dataset and compare it with a prompting baseline. The graded core is a correct, controlled experiment and an evidence-based interpretation of the training behavior and evaluation results.

HW2. Preference data and DPO (released Oct 21, due Nov 4, 15%)

Use a provided pool of prompts and candidate responses to make and document your own preference judgments, then run a controlled DPO experiment and compare it with a reference checkpoint. The emphasis is on human labeling, controlled evaluation, and interpretation of behavioral changes and possible regressions.

Concept check (10%, in class Nov 11)

A short, handwritten, closed-device assessment covering the core concepts of the first eleven weeks. It exists to confirm the core ideas landed.

Group project (45%, teams of 3)

A semester-long project in which your team chooses a real task, builds a credible baseline and evaluation suite, and conducts a scoped post-training experiment. A negative result that is well evaluated and carefully argued can still be a successful project. The final report should incorporate feedback received during the project presentation.

Milestone	Weight	Due	Deliverable
Proposal	5%	Sep 23	One-page project brief covering the task rationale, baseline and data plan, and initial evaluation design.
Project checkpoint	5%	Oct 28	5 minute oral progress update followed by two minutes of questions supported by one structured checkpoint slide.
Final presentation	10%	Nov 18	In-class presentation with feedback for final revision
Revised final report and code	25%	Dec 7	Written report incorporating presentation feedback and a reproducible code repository

Grading Policy

Component	Weight	Released	Due
Participation and in-class labs	10%		
HW0: TACC inference and resource accounting	5%	Aug 26	Sep 9
HW1: Supervised fine-tuning and LoRA	15%	Sep 23	Oct 7
HW2: Preference data and DPO	15%	Oct 21	Nov 4
Concept check (in class, handwritten)	10%		Nov 11
Group project proposal	5%		Sep 23
Group project checkpoint	5%		Oct 28
Group project final presentation	10%		Nov 18
Group project revised final report and code	25%		Dec 7

Plus and minus grades will be used in the final class grade. Final percentages are not rounded up.

Attendance and Participation

Because we meet only once a week and much of the lab work happens collaboratively in the room, attendance matters. The 10% participation component reflects attendance and engagement in labs. You may miss up to two class meetings without affecting your participation grade. After those two, each additional unexcused absence lowers the final course grade by two percentage points through the participation component.

Absences for illness, religious holidays, university-sponsored activities, or documented emergencies are excused. Excused absences do not count against your participation grade or your two permitted absences. If you miss a lab, the notebook and instructions will be on Canvas, and the TA can help you catch up in office hours.

Late Work and Makeup Policy

Each student has **four late days** for the semester, usable in whole-day increments across HW0, HW1, and HW2, with no penalty and no explanation required. Beyond that, late work loses 10% per day. Project milestones are not eligible for late days.

Final Examination

There is no final examination in this course. No examination will be scheduled during the University final examination period. Final project presentations take place on November 18, 2026. The December 2 class is

devoted to feedback-based project revision and refinement, and the revised final project report and code are due at 11:59 pm on December 9, 2026.

Course Artificial Intelligence Policy

In accordance with the University's Institutional Rules on Student Services and Activities, Chapter 11, students accept the responsibility to always uphold academic integrity and an honor code reflective of a scholarly community devoted to academic and personal success. All members of the University community are fully accountable and responsible for any output they produce as part of academic work. They are also responsible for following the guidance specified in the Texas Statement on Academic Integrity and avoiding prohibited uses of generative AI tools outlined in the Information Security Office's guidance on Acceptable Use of Generative AI Tools.

In this course, generative AI use is partially permitted.

This is a course about language models and I expect you to use them. By default, generative AI tools are allowed for coursework as long as you review their output critically, remain responsible for what you submit, and disclose use that contributed to the submitted work. However, there are deliberate exceptions (listed below) to cover the parts of the course where your own judgment is what I am trying to teach and assess. Any assignment-specific restrictions will be stated in writing.

Activity	AI tools	Notes
Debugging, environment setup, boilerplate training/eval scripts	Permitted	Disclose in your AI use statement
Explaining concepts, papers, or error messages to yourself	Permitted	No disclosure required
Drafting, editing, and polishing prose in reports	Permitted	Disclose; you own every claim
Generating synthetic training data where the assignment calls for it	Permitted	Describe prompts and filtering in your write-up
Choosing your project task and domain	Not permitted	Your team must choose the task and domain yourselves. AI may be used in the brainstorming stage, but the final decision needs to be made by humans.
Authoring evaluation prompts, rubrics, and gold labels	Not permitted	You or your team write these; AI may critique a draft you wrote
Selecting and labeling preference pairs (HW2)	Not permitted	Human judgment is the point of the assignment
Interpreting your training runs and writing the analysis	Not permitted	The interpretation and conclusions must be yours
Concept check	Not permitted	Closed-device, handwritten, in class

If a situation is not covered above, please feel free to ask me.

Disclosure Policy

Every graded submission must include a short AI Use Statement at the end of the document. Describe AI use that contributed directly to the submitted work. Routine studying or concept clarification does not need to be listed. The statement should include:

- Which tools you used (e.g., Claude, ChatGPT, GitHub Copilot, Cursor, Codex)
- What you used them for, by task (e.g., “debugging the SLURM submission script”)
- Which parts of the submission were substantially AI-generated and how you verified them

Two or three sentences is usually enough. If you used no AI tools in preparing the submission, say so. For group submissions, each member writes a separate statement.

Failing to disclose AI use when this policy requires disclosure may constitute academic misconduct under UT Austin's Institutional Rules and may be referred to Student Conduct and Academic Integrity in the Office of the Dean of Students.

Do not upload other students' work, unpublished course materials, or any non-public data to external AI services.

University Policies

Academic Integrity

All allegations of academic misconduct will be referred by faculty to the Office of the Dean of Students. Students who violate University rules on academic misconduct are subject to the student conduct process. A student found responsible for academic misconduct may be assigned both a status sanction and a grade impact for the course. The grade impact could range from a zero on the assignment in question up to a failing grade in the course. A status sanction can include a written warning, probation, deferred suspension, suspension, or expulsion from the University. To learn more, visit deanofstudents.utexas.edu/conduct/student-conduct-processes.php

Collaboration

Homework is individual work. You may discuss approaches, debug together, and share environment tips freely. However, the code, data, and write-up you submit must be your own. The project is group work; all members are expected to contribute substantively, and the final report includes a short statement of individual contributions.

Classroom Recording (HOP 2-9970)

HOP 2-9970 prohibits students from recording class instruction (audio or video) unless a student obtains the instructor's permission or Disability & Access has approved audio recording as an accommodation.

Sharing of Course Materials Is Prohibited

No materials used in this class, including, but not limited to, lecture hand-outs, videos, assessments (quizzes, exams, papers, projects, homework assignments), in-class materials, review sheets and additional problem sets, may be shared online or with anyone outside of the class without my explicit, written permission.

Disability and Access (D&A)

The university is committed to creating an accessible and inclusive learning environment consistent with university policy and federal and state law. Please let me know if you experience any barriers to learning so I can work with you to ensure you have equal opportunity to participate fully in this course. If you are a student with a disability, or think you may have a disability, and need accommodations please contact Disability & Access (D&A). Please refer to the D&A website for more information: disability.utexas.edu

If you are already registered with D&A, please deliver your Accommodation Letter to me as early as possible in the semester so we can discuss your approved accommodations and needs in this course.

Religious Holy Days

By UT Austin policy, you must notify me of your pending absence for a religious holy day as far in advance as possible of the date of observance. If you must miss a class, an examination, a work assignment, or a project to observe a religious holy day, please fill out the required form administered by the Office of the Dean of Students. You will be given an opportunity to complete the missed work within a reasonable time after the absence, as per Texas Education Code Section 51.911.

Names and Pronouns

Class rosters are provided to the instructor with the student's legal name, unless they have added a chosen name with the registrar's office. I will gladly honor your request to address you with the name and pronouns you prefer. Please advise me of any changes or needs regarding your name and pronouns at any time.

Counseling and Mental Health

Students who are struggling for any reason and who believe it might affect their performance in the course are encouraged to reach out. The Counseling and Mental Health Center (CMHC) can be reached at 512-471-3515, and the CMHC 24/7 Crisis Line at 512-471-2255.

Title IX Disclosure

Beginning January 1, 2020, Texas Education Code, Section 51.252 (formerly known as Senate Bill 212) requires all employees of Texas universities, including faculty, to report any information to the Title IX Office regarding sexual harassment, sexual assault, dating violence and stalking that is disclosed to them. Texas law requires that all employees who witness or receive any information of this type (including, but not limited to, writing assignments, class discussions, or one-on-one conversations) must be reported. I am a Responsible Employee and must report any Title IX related incidents that are disclosed in writing, discussion, or one-on-one. Before talking with me, or with any faculty or staff member about a Title IX related incident, be sure to ask whether they are a responsible employee.

Carrying of Handguns on Campus

Students in this class who hold a license to carry are asked to review the university policy regarding campus carry. Open carry is not permitted anywhere on campus. I prohibit the carrying of handguns in my personal office.

Campus Safety

Students should sign up for Longhorn Alert emergency text notifications. Occupants of campus buildings must evacuate when a fire alarm is activated; please familiarize yourself with the exits of our classroom and follow the instruction of faculty during any evacuation. Students requiring assistance in evacuation should inform me in writing during the first week of class.

University Policies and Resources for Students

A list of resources important for engaging with this course and the university is available on the University Policies and Resources for Students Canvas page: <https://utexas.instructure.com/courses/1240481>

Land Acknowledgment

I would like to acknowledge that we are meeting on the Indigenous lands of Turtle Island, the ancestral name for what now is called North America. I would like to acknowledge the Alabama-Coushatta, Caddo, Carrizo/Comecrudo, Coahuiltecan, Comanche, Kickapoo, Lipan Apache, Tonkawa and Ysleta Del Sur Pueblo, and all the American Indian and Indigenous Peoples and communities who have been or have become a part of these lands and territories in Texas.