UNIVERSITY OF WATERLOO • FALL 2026

CS 886: Learning Theory for Modern AI

Transformers and Large Language Models

Kimon Fountoulakis
Associate Professor, Cheriton School of Computer Science

watcl-lab.github.io/cs886-learning-theory/

What this course studies

COURSE SCOPE

- **Representation:** what functions, languages, and algorithms can transformers express?
- **Computation:** what depth, precision, communication, and running time do they require?
- **Learning:** when is training stable, and what structures does gradient-based training discover?
- **Generalization and reasoning:** how do in-context learning and chain-of-thought behave on new tasks and longer inputs?

Course roadmap

TWELVE MEETINGS

Week 1	**Opening case study**	Exact algorithmic learning, certification, and hardness. Instructor-led.
Weeks 2–5	**Module I**	Expressivity, computational complexity, trainability, and learning sparse attention.
Weeks 6–8	**Module II**	Statistical targets, optimization, and finite-sample generalization in in-context learning.
Weeks 9–10	**Module III**	Autoregressive chain-of-thought, curricula, scratchpads, and length generalization.
Weeks 11–12	**Module IV**	Individual project presentations.

Course logistics

MEETING DETAILS

Meeting time	Fridays, 1:30–4:20 p.m.
Delivery	In person.
Seminar meetings	September 11–December 4, 2026.
Reading Week	No class on October 16; Reading Week is October 10–18.
Office hours	Email the instructor at kimon.fountoulakis@uwaterloo.ca.
Website	watcl-lab.github.io/cs886-learning-theory/

Student paper presentations

ASSIGNMENT STRUCTURE

Enrollment	27 students.
Paper slots	27 joint paper-presentation slots (54 presenter assignments) across Weeks 2–10.
Weekly schedule	Three papers per week, each jointly presented by two students.
Student workload	Each student gives two paired paper presentations on different papers and will normally work with a different partner for the two talks.
Week 1	The instructor presents all four opening-case-study papers.

The individual project presentation is separate and does not count toward the two required paired paper presentations.

Presentation timing

STUDENT-LED MEETINGS: WEEKS 2–10

44 min	**Joint presentation**	One coherent talk by two students; each normally presents about 22 minutes of substantive technical material.
10 min	**Discussion**	Questions about assumptions, proof details, limitations, and interpretation.
1 min	**Transition**	Move promptly to the next paper and presenting pair.
	Meeting order	Papers 1–2 (110 minutes), a 5-minute break, then Paper 3 (55 minutes).
170 min	**Total time**	Three 55-minute paper slots plus the break.

Paired presentation responsibilities

SHARED PREPARATION, INDIVIDUAL EVALUATION

- Each pair prepares one coherent 44-minute presentation, not two separate mini-presentations.
- Speaking time and technical content should be divided approximately equally; each student normally presents about 22 minutes and a substantive technical part of the paper.
- Both presenters must understand the complete paper, including its formal problem, main theorem, assumptions, proof strategy, and limitations.
- Both presenters should be ready to answer general questions about any part of the paper.
- Each student receives an individual grade. Organization and coherence may be assessed jointly; technical understanding, delivery, contribution, and responses to questions are assessed individually.

Presentation content

WHAT TO EXPLAIN

- The **formal problem**: data-generating process, architecture or hypothesis class, loss, training rule, and evaluation criterion
- The **claim type**: representation, computational complexity, trainability, optimization, learnability, generalization, or certification
- The **main theorem**: assumptions, quantifiers, and important parameter dependence
- The **proof mechanism**: the central argument, not only the conclusion

Assessment

FINAL GRADE

40%	**Paired paper presentations**	Two paired presentations on different papers, worth 20% each; each student receives an individual grade for each.
30%	**Project report**	Final written report and required reproducibility materials.
10%	**Project presentation**	Individual talk in Week 11 or Week 12.
20%	**Participation**	Class participation.

Course project: 40% total = 30% report + 10% presentation.

Two project paths

CHOOSE A RESEARCH CONTRIBUTION

New theoretical contribution

- A new theorem, proof, or sharper bound
- Weaker assumptions or a broader model
- An impossibility result or counterexample

Empirical validation or scrutiny of theory

- Test a formal prediction in finite models
- Relax an idealized assumption
- Identify failure regimes or discrepancies

Previous project. A project from a previous offering of CS 886 was subsequently developed into a NeurIPS 2024 publication.

What makes a strong project

PROJECT STANDARD

- A precise research question tied to transformer or language-model theory
- A technically substantive original contribution, including careful negative results
- Clear separation between prior work and the student's contribution
- Reproducible code and configurations for empirical projects

A generic model comparison, benchmark leaderboard, prompt-engineering exercise, or systems implementation is not sufficient by itself.

Use of generative AI

ACADEMIC RESPONSIBILITY

- Generative-AI tools may be used for coursework in this class.
- Students remain fully responsible for the accuracy, originality, and integrity of all submitted work.
- Students must be able to explain every mathematical statement, proof step, citation, experimental result, and piece of code.
- Factual, mathematical, citation, experimental, or coding errors are graded as errors regardless of whether AI was used.

Project timeline

THREE MILESTONES

Oct. 23	**Topic approval**	Ungraded three-to-five-sentence description of the contribution and, for empirical work, the principal experiment.
Nov. 26	**Final report**	Written report and required reproducibility materials; 30% of the final grade.
Nov. 27 / Dec. 4	**Project presentation**	Individual presentation on the assigned date; 10% of the final grade.

All dates are in 2026. Assigned dates will be announced after topic confirmation.

Project presentations

WEEKS 11–12

Total presentations	27 individual project talks.
Week 11	14 presentations on November 27, 2026.
Week 12	13 presentations on December 4, 2026.
Length	Approximately 12–13 minutes per talk, including questions.
Assessment	10% of the final grade, within the 40% course project.

The project presentation is separate from the two required paired paper presentations.

Course materials and website

WHERE TO FIND EVERYTHING

Textbook	No textbook is required.
Readings	All assigned papers are linked from the course website.
Website	watcl-lab.github.io/cs886-learning-theory/

Contact

QUESTIONS ABOUT THE COURSE

Kimon Fountoulakis

kimon.fountoulakis@uwaterloo.ca

Office: DC 3611

watcl-lab.github.io/cs886-learning-theory/