WISCONSIN
UNIVERSITY OF WISCONSIN–MADISON

CS 839 Theoretical Foundations of Deep Learning, Spring 2022 Syllabus

Key Course Offering Information

General Identifying Information

Institution Name: University of Wisconsin–Madison
Course Subject, Number and Title: CS 839 Theoretical Foundations of Deep Learning
Credits: 3 credits
Course Designations and Attributes: Grad 50% - Counts toward 50% graduate coursework requirement
Course Description: Topics selected from advanced areas.
Requisites: Graduate/professional standing.
Meeting Time and Location: Tu/Th 11AM – 12:15PM, CS 1325
Instructional Modality: In-person
Instructor Contact Info: Yingyu Liang, CS5387, office hour Thursday 9:45-10:45am and 1-2pm, yliang@cs.wisc.edu
Teaching Assistant Contact Info (if applicable): None

Course Learning Outcomes

1. Identify and summarize the key challenges in theoretical analysis of deep learning.
2. Identify and summarize the approximation power of neural networks, including approximating Liptschitz-continuous functions, universal approximation, and Barron's theorem.
3. Identify the implicit regularization in linear regression, logistic regression, and training neural networks. Demonstrate the knowledge about Clarke Subdifferential and positive homogeneity.
4. Identify the Neural Tangent Kernel view. Demonstrate the knowledge about its different formulations.
5. Summarize and demonstrate the knowledge about the mean-field theory for two-layer neural networks.
6. Illustrate the representation learning paradigm and foundations models. Identify the analysis of the principle for general representation learning and contrastive learning.
7. Demonstrate the knowledge about the computation lower bounds for network learning, including NP-hardness and Statistical Query models.
8. Identify and present key concepts in chosen papers in the provided reading list.
9. Identify research questions and perform research on the chosen question.

How Credit Hours are Met by the Course

This class meets for two, 75-minute class periods each week over the spring semester and carries the expectation that students will work on course learning activities (reading, writing, problem sets, studying, etc) for about 3 hours out of the classroom for every class period.

Regular and Substantive Student-Instructor Interaction

Interaction with faculty and instructional staff in this course include:

1. *Participation in regularly scheduled lectures*
2. *Personalized comments for each student's assignment and exam*
3. *Interaction with students on Canvas and Piazza, including answering questions, making announcements, setting up polls, etc.*
4. *Weekly office hours*

Instructor-to-Student Communication

Course Overview

Deep learning has been the main driving force behind many modern intelligent systems and has achieved great success in many applications such as image processing, speech recognition, and game playing. However, the fundamental questions about why deep learning is so successful remain largely open. The goal of this course is to study and build the theoretical foundations of deep learning. Topics covered by this course include but are not limited to: approximation power of neural networks, optimization for deep learning, generalization analysis of deep learning. The instructor will give lectures on the selected topics. Students will present and discuss papers on the reading list, and do a course project.

The course will consist of mostly reading and discussing recent important papers on the theoretical analysis of deep learning, some homework assignments, and a course project.

Course Website and Digital Instructional Tools

- Course website: https://pages.cs.wisc.edu/~yliang/cs839_spring22/index.html
- Canvas: https://canvas.wisc.edu/courses/295404

Discussion and/or Laboratory Sessions

None.

Required Textbook, Software and Other Course Materials

- No required textbook.
- The students are expected to be familiar with the analysis tools in the following textbooks (or at similar levels):
  - Understanding machine learning: From theory to algorithms. Shai Shalev-Shwartz, and Shai Ben-David. Cambridge University Press, 2014.
  - High-Dimensional Probability: An Introduction with Applications in Data Science. Roman Vershynin. Cambridge University Press, 2018.
  - Introductory Lectures on Convex Optimization: A Basic Course. Yurii Nesterov. Springer, 2004.

Homework and Other Assignments

- About 5 homework assignments.
- Assignments will be posted and submitted on Canvas.

Homework is required to be written in Latex. Unless indicated otherwise, you can discuss with the other students but must finish the homework by yourself. If you discuss with others, please indicate that in your submission; if you consult external materials like Internet post, please cite the references.

Exams, Quizzes, Papers and Other Major Graded Work

- There will be no exams but a course project.
- Students are required to do a project in this class, since the goal of the course is to provide the opportunity to explore the frontier in recent theoretical studies of deep learning. A project guideline will be provided to specify the details. Roughly speaking, projects should be proposed by the proposal deadline (this is expected to be around the midterm and will be specified in class). A pdf report (written in Latex) should be submitted by the project deadline. The report should be in the style of a conference paper, providing an introduction/motivation, discussion of related work, a description of your work that is detailed enough that the work could be replicated, and a conclusion.
- The topic of the project include but not limited to:
    - Extension of existing work: improved bounds, more thorough analysis, adaptation to new problem settings, etc
    - Novel theoretical analysis of existing deep learning methods/problems
    - Novel formulation of existing deep learning problems and corresponding analysis
    - Interesting empirical observations and proposing theoretical explanations (preferably in the form of math analysis)
- The ideal outcome is a report publishable in major machine learning or theory conferences or journals. Published work of the students cannot be used as the course project.

Course Schedule/Calendar

See the course website: https://pages.cs.wisc.edu/~yliang/cs839_spring22/schedule.html

Grading

The grading for the course will be based on (temporally, may subject to changes later):

- Lecture note scribe: 10%
- Homework assignments: 40%
- Paper presentation: 10%
- Course Project: 40%

Academic Policies and Statements

ACADEMIC INTEGRITY STATEMENT

By virtue of enrollment, each student agrees to uphold the high academic standards of the University of Wisconsin-Madison; academic misconduct is behavior that negatively impacts the integrity of the institution. Cheating, fabrication, plagiarism, unauthorized collaboration, and helping others commit these previously listed acts are examples of misconduct which may result in disciplinary action. Examples of disciplinary sanctions include, but are not limited to, failure on the assignment/course, written reprimand, disciplinary probation, suspension, or expulsion.

DIVERSITY & INCLUSION STATEMENT

Diversity is a source of strength, creativity, and innovation for UW-Madison. We value the contributions of each person and respect the profound ways their identity, culture, background, experience, status, abilities, and opinion enrich the university community. We commit ourselves to the pursuit of excellence in teaching, research, outreach, and diversity as inextricably linked goals. The University of Wisconsin-Madison fulfills its public mission by creating a welcoming and inclusive community for people from every background – people who as students, faculty, and staff serve Wisconsin and the world.