CS 577 — Deep Learning — Syllabus

Fall 2024

Instructor:	Yutong Wang	**Time:**	W 17:00 – 19:40
Email:	ywang562@iit.edu	**Place:**	Wishnick Hall 113.

Course Overview: Deep neural networks form an important sub-field of machine learning that is responsible for much of the progress in cognitive computing in recent years in areas of computer vision, audio processing, and natural language processing. Deep networks can be trained with a single end-to-end model and bypass the need for traditional task-specific feature engineering. In this way deep learning simplifies learning tasks and allows using developed models to new tasks. Deep networks are suitable for parallel processing implementations and can easily leverage intensive computational resources. The course will focus on mathematical concepts, numerical algorithms, principles, GPU frameworks, and applications of deep learning. Topics include deep feedforward networks, convolutional networks, sequence modeling, transformers, and deep generative models with applications to data analysis, computer vision, and natural language processing. Several programming assignments and a project will practice the application of deep learning techniques to actual problems. The course requires sufficient math and programming background but does not require prior knowledge in machine learning.

Learning objectives: By taking this course, students will

- gain experience with designing a simplified deep learning framework from scratch,
- understand the internals of deep learning frameworks such as PyTorch,
- develop mathematical skills for further study in theoretical principles and concrete implementations of state-of-the-art deep learning models,
- develop analytical skills to be able to efficiently and critically read the latest research,
- develop writing and communication skills to explain complex concepts via the course project.

Course at a glance: There will be 14 lectures, 7 homeworks, 1 (in class) midterm, 1 peer-graded group project, and 1 final exam.

Important Dates:

- Aug 21 - First lecture
- Oct 2 - Midterm exam
- Oct 9 - DUE: Course project group formation and paper selection (see “Course Project” section below)
- Nov 15 - DUE: Course project
- Nov 16 - Reviewer-to-project assignment is released
- Nov 25 - DUE: Course project peer grading
- Nov 20 - Last lecture
- Dec 2-7 - Final exam (exact date TBD)

Prerequisites: CS 430 is the only prerequisite course. In addition, student should be

- comfortable programming in Python, working with matrices in NumPy/Julia/MATLAB, and using Jupyter Notebook,
- comfortable with vector calculus, although we will briefly review the necessary material,
- willing to learn LaTeX (see the "LaTeX and Homework" section below).

Note:

- You do NOT need to have worked with PyTorch, although experience will be helpful.
- This course will not require access to GPUs.

Lecture structure and contents: There are 14 lectures in total. Each lecture is 2 hours and 40 minutes long. There is a 15-minute break in the middle of every lecture. No attendance will be taken. No work done during lecture will be graded (other than the midterm and final exam). Any in-class work are purely for additional practice. However, you are highly encouraged to attend lectures, participate, and ask questions as much as possible.

- Lec 01 - Aug 21 - Backgrounds on linear algebra, calculus, python, LaTeX, and machine learning
- Lec 02 - Aug 28 - Regression: from linear regression to kernel methods to neural networks
- Lec 03 - Sep 04 - Classification: loss functions, Cross entropy, numerical stability
- Lec 04 - Sep 11 - Optimizers: momentum, adaptive methods
- Lec 05 - Sep 18 - Backpropagation part I
- Lec 06 - Sep 25 - Backpropagation part II
- Lec 07 - Oct 02 - 1st half: Midterm exam. 2nd half: Computer vision, part I: convolutional networks
- Lec 08 - Oct 09 - Computer vision, part II: diffusion models
- Lec 09 - Oct 16 - Natural language processing, part I: recurrent networks
- Lec 10 - Oct 23 - Natural language processing, part II: transformers
- Lec 11 - Oct 30 - Hardware, part I: computation CPU versus GPU
- Lec 12 - Nov 06 - Hardware, part II: Flash-Attention
- Lec 13 - Nov 13 - Theory: implicit regularization
- Lec 14 - Nov 20 - Review for final exam

The contents for the lectures are tentative. I will slightly adjust as the semester progresses.

Homeworks: There will be 7 homeworks. The final homework, HW 7, is for participation in the peer review for the course project. See the "Course Project" section below. HW 1 to HW 6 are "regular" homeworks, which consisting of a math part and a programming part.

- The math part of the homework must be turned in as a single PDF file generated by LaTeX using the template we provide.
- The programming part of the homework must be turned in as a python notebook file (i.e., a `.ipynb` file).

- We will not run your code, so please double check that all outputs are visible.

Homework and LaTeX: LaTeX is a tool for creating academic documents containing mathematics. All homework must be submitted as PDFs generated by LaTeX. To help you get started, parts of the first homework assignment (HW 1) will be a LaTeX tutorial. We will provide templates for you to use.

Note:

- Overleaf (`https://overleaf.com/`) allows you to use LaTeX on the web without installing LaTeX on your own computer.
- Search the internet for "online latex equation editor" to practice typesetting individuals equations.

Homework due dates:

- HW 1 - Due Aug 27 - Basics of LaTeX and NumPy/PyTorch
- HW 2 - Due Sep 10 - "Shallow" learning methods
- HW 3 - Due Sep 24 - Optimizers and backpropagation
- HW 4 - Due Oct 8 - More backpropagation and CNNs
- HW 5 - Due Oct 22 - Diffusion models and RNNs
- HW 6 - Due Nov 5 - Transformers, GPUs
- HW 7 - Due Nov 25 - Project peer review (See "Course Project — Peer Grading" section for detail)

Homework is due on the due date at 11:59pm Central Time.
HW 1 will be released on Aug 21 right before the Lecture 01.
HW 2-6 will be released on the day after the previous homework's due date. For example, HW 1 is due on Aug 27, and HW 2 will be released on Aug 28 before the Lecture 02.
HW 7 will be "released" when Nov 16 when the reviewer-to-project assignment is released.

Course Project: In the project, you will turn in a single python notebook file (`.ipynb`) that "annotates" a deep learning model from your choice of a published paper in one of the venues in the "Venues for Paper Selection" section. The goal is to create a detailed and well-annotated explanation of the model, similar in style to "The Annotated Transformer" blog post (`https://nlp.seas.harvard.edu/annotated-transformer/`).

- The project will be due at 11:59pm Central Time on Nov 15.
- You must work in groups of size 2, 3 or 4 students.
- You must finalize your group roster and paper choice by emailing me and all TAs before 11:59pm Central Time on Oct 9.

Course Project — requirements: Your project must include each of the 6 parts below:

1. Background:
   - Explain clearly the motivation: the problem does the paper seek to address
   - Explain the key breakthrough of the paper: which innovations enabled the paper to resolve the problem

2. Model architecture:
   - Analysis: Use diagrams/visualizations of the model together with code to explain the model, part-by-part
   - Synthesis: Use a diagram/visualization of the model together with code to explain how the parts come together
3. Model training, explain the training loop and its sub-parts. For instances, in "The Annotated Transformer", the authors of the blog discussed "Optimizer" and "Regularization", among several other items. There is no standard set of parts to a training loop. But be as thorough as possible in explaining all the training techniques that are used.
4. Minimal CPU-ready working example (MWE-CPU).
   - Create a small synthetic dataset OR find a small real world dataset suitable for the model
   - Create a miniaturized version of the model that trains on a laptop. You can achieve this by reducing the number of parameters/layers/width/floating point precision/or all of the above. "Everything should be made as simple as possible, but no simpler."
   - Train the miniaturized model on the small dataset and discuss the training process and outcome. Note: our grading criteria is not with respect to the model's performance. Instead, we are looking for maximizing our understanding.
5. Discussion: Weaknesses/limitations/future directions.
6. Group member contribution: Detail the contributions of each group member.

Course Project — Peer Grading: Every member of the class is required to review 4 projects. The reviewer-to-project assignment will be done randomly and released on Nov 16 by noon. For the review, you will be asked to read each assigned project carefully and assign scores based on several criteria that will be provided later. In addition to the scores, you will be asked to turn in a plain text file justifying your reason for each score. The review will be double blind (either the reviewers nor the authors will know the identity of each other).

If the project receives peer review scores that has a reasonable level of variance, the overall project grade will be the average review score. In case that a project receives unusually high variance in the scores, the instructors will examine the project and reviewers justifications closely and intervene if necessary.

Note on academic integrity and peer review: It is possible that the reviewer may learn the identity of the authors unintentionally (e.g., via conversations prior to Nov 16 the reviewer-to-project assignment released). This is allowed and NOT considered a breach of academic integrity.

However, after Nov 16, reviewers must refrain from learning the identity of the authors. Purposely circumventing anonymity to gain advantage *is* considered academic dishonesty. See section on "Academic Integrity" below.

Course Project — Venues for Paper Selection:

AAAI - Association for the Advancement of Artificial Intelligence Conference, ACL - Annual Meeting of the Association for Computational Linguistics, CVPR - Conference on Computer Vision and Pattern Recognition, ECCV - European Conference on Computer Vision, EMNLP - Conference on Empirical Methods in Natural Language Processing, ICCV - International Conference on Computer Vision, ICLR - International Conference on Learning Representations, ICML - International Conference on Machine Learning, IJCAI - International Joint Conference on Artificial Intelligence, KDD - Conference on Knowledge Discovery and Data Mining, NAACL - North American Chapter of the Association for Computational Linguistics Conference, NeurIPS - Conference on Neural Information Processing Systems, SIGIR - SIGIR Conference on Research and Development in Information Retrieval, WWW - International World Wide Web Conference.

Exams: The midterm will cover material from Lectures 1 through 6. The final exam will cover all material, with emphasis on material from Lecture 7 and onward. Both exams will be a written exam.

On each exam, you may bring a "cheat-sheet" front and back on a single sheet of paper. The paper must be equal to or smaller than (i.e., fit inside) either 8.5 x 11 in (US standard) OR 210 x 297 mm (International standard).

If you need to reschedule either exam for any reason, please contact me as far in advance as possible so we can make suitable arrangements. Once you sit for an exam, the score you earn on it is final (i.e., there are no "retakes").

Grading: Your final grade will be computed as follows:
30% Homeworks (Each homework is worth 5%. Lowest grade on HW 1 through 6 will be dropped.)
30% Course Project
20% Midterm exam
20% Final exam

- Course grade scale $\{A > 87\%; B > 74\%; C > 61\%; D > 48\%; E \leq 48\%\}$.
- I might adjust the scale by $\pm 2\%$.
- Exam scores may be linearly scaled so that the class median/average (whichever is lower) is 74%. I will publish the scaling formula if and when I do this.

Textbook: There is no required textbook for the course. However, you may find the following textbooks helpful:

Goodfellow, Ian, Yoshua Bengio, and Aaron Courville. *Deep Learning.* MIT Press, 2016. https://www.deeplearningbook.org/

Zhang, Aston, Zachary C. Lipton, Mu Li, and Alexander J. Smola. *Dive into Deep Learning.* Cambridge University Press, 2023. https://d2l.ai/

Late Policy: Homeworks turned in 1 day late will receive a 33% penalty. Homeworks turned in 2 days late will receive a 66% penalty. Homeworks turned in 3 or more days late will receive a zero. Lowest grade on HW 1 through 6 will be dropped.

Collaboration: You are welcome to discuss HW 1 through 6 with classmates, but all final work must be your own. You may not discuss the peer review with other students (HW 7).

Academic Integrity:* Academic dishonesty of any kind may result in a 0 on the homework/project, a reduction in final grade, and/or referral to the Dean.

The IIT code of Academic Honesty may be found in the handbook.
https://www.iit.edu/student-affairs/student-handbook/fine-print/code-conduct

Disability Accommodations: Reasonable accommodations will be made for students with documented disabilities. In order to receive accommodations, students must obtain a letter of accommodation from the Center for Disability Resources. The Center for Disability Resources (CDR) is located in Life Sciences Room 218, telephone 312 567.5744 or disabilities@iit.edu.

Sexual Harassment and Discrimination Information: Illinois Tech prohibits all sexual harassment, sexual misconduct, and sex discrimination by any member of our community. This includes harassment among students, staff, or faculty. Sexual harassment of a student by a faculty member or sexual harassment of an employee by a supervisor is particularly serious. Such conduct may easily create an intimidating, hostile, or offensive environment.

Illinois Tech encourages anyone experiencing sexual harassment or sexual misconduct to speak with the Office of Title IX Compliance for information on support options and the resolution process.

You can report sexual harassment electronically at iit.edu/incidentreport, which may be completed anonymously. You may additionally report by contacting the Title IX Coordinator, Virginia Foster at foster@iit.edu or the Deputy Title IX Coordinator, Cat Rodl at student.health@iit.edu. Your voice matters, and we are here to help.

For confidential support, you may reach Illinois Tech's Confidential Advisor at (773)907-1062. You can also contact a licensed practitioner in Illinois Tech's Student Health and Wellness Center at student.health@iit.edu or (312)567-7550

For a comprehensive list of resources regarding counseling services, medical assistance, legal assistance and visa and immigration services, you can visit the Office of Title IX Compliance website at https://www.iit.edu/title-ix/resources.