Computer Engineering Machine Learning and Deep Neural Nets
ECE 590-10/11

Fall 2019
Syllabus

Course Summary

This course examines various computer engineering methods commonly performed in developing machine learning and deep neural network models. The focus of the course is how to improve the training and inference performance in term of model accuracy, size, runtime, etc. Techniques that are widely investigated and adopted in industrial companies and academic communities will be discussed and practiced. Programming practices on these techniques are designed with heavy utilization of PyTorch package. Classes will be running by combining lecture sessions and discussion sessions for the programming practices.

Learning Objectives

This course is designed to improve your ability to:

1. **Comprehend** the mechanisms, applications, and limitations of techniques commonly used in training and inference of machine learning and deep neural networks algorithms;
2. **Formulate** hypotheses and **conduct** experiments employing these techniques;
3. **Analyze** experimental results obtained by these techniques and your own practices and **derive** the conclusions that are supported or not supported by your data;
4. **Synthesize** and **communicate** the experimental results and data through oral narrative, graphs, figure legends, and result narratives;
5. **Utilize** proper engineering techniques for novel machine learning algorithms and deep neural network models;
6. **Propose** new engineering approaches and techniques to further enhance machine learning and deep neural network training and inference execution.

Logistics

	ECE 590-10	**ECE 590-11**
Course Faculty:	Dr. Hai "Helen" Li Room 209B Hudson Hall Hai.li@duke.edu	Dr. Yiran Chen Room 209B Hudson Hall Yiran.chen@duke.edu
Lectures:	Tuesday/Thursday 10:05-11:20am Hudson Hall 208	Tuesday/Thursday 1:25-2:40pm Hudson Hall 207
Office Hours:	By appointment	By appointment
Teaching Assistants:	Tunhou Zhang tunhou.zhang@duke.edu	Huanrui Yang huanrui.yang@duke.edu
	Qing Yang qing.yang21@duke.edu	Meng Xia mx41@duke.edu
Prerequisite:	We expect that students to have basic object-oriented programming experience (e.g. C++, Python) and be familiar with linear algebra and computer hardware fundamentals prior to taking this course, such as • For graduate students: ECE 550 + ECE 551 • For undergraduate students: ECE 381 + CS 308 + ECE/CS 250.	

If you do not have these pre-requisites and are unfamiliar with these topics, please note that we will not be slowing down to cover them. We apologize about this. But we must cover a respectable portion of the course material and there are a limited number of lecture sessions.

Textbooks: There are no mandatary textbooks for this course.
The related reading materials (e.g., papers, webpages, etc.) will be distributed through Sakai before the classes.

Software: We recommend downloading Pytorch (https://pytorch.org/)

Grading: Lab assignments (4): 65%
Homework assignments (3): 30%
In-class assignments/discussion: 5%

Completion of all assignments is required in order to earn a passing grade of D- or better in this course.

Course Grade: The weighting of the graded components is provided above.
Course grades are determined using an absolute, but adjustable scale (i.e., there is no curve). A final course average (rounded to the nearest 0.1 point) of at least 93.3 = A, 90.0 = A-, 86.7 = B+, 83.3 = B, 80.0 = B-, etc.

Note: the professors reserve the rights to scale the grades.

Topics to be Covered and Tentative Schedule (Subjective to Change)

Week	Date	Topic
1	8/27	Course Introduction
	8/29	Deep Learning Introduction
2	9/3	NumPy/Pytorch Tutorial
	9/5	Convolutional Neural Network (CNN)
3	9/10	Forward-Backward Propagation
	9/12	CNN Training
4	9/17	CNN Training (cont'd)
	9/19	CNN Architecture
5	9/24	Compact Neural Architecture Design
	9/26	RNN and Language Models
6	10/1	Distributed Computing
	10/3	Hardware Systems - GPUs, CPUs, Cloud Servers
7	10/8	*Fall Break. NO Class.*
	10/10	Model Compression I
8	10/15	Model Compression II
	10/17	Pruning & Quantization: LAB + TAs
9	10/22	Sparse Regularization
	10/24	Neural Architecture Search
10	10/29	Sparse Optimization
	10/31	DNN Quantization
11	11/5	Hardware Architecture - Accelerators, RRAM, etc.
	11/7	General discussion on machine learning security
12	11/12	Adversarial Attacks and Defense I
	11/14	Adversarial Attacks and Defense II
13	11/19	Generative Adversarial Network
	11/21	Decentralized and Privacy
14	11/26	Lab + TAs

Course Procedure

Communication

We will use **Sakai** announcements to communicate with you outside of class. Typical announcements include posting of course materials and changes of office hours.

By default, when I post an announcement, I will also send a "low priority" email. You have the option as to whether or not to receive them. My understanding is that by default you will receive low priority emails. To view or change your notification options go to "Home" then select "Preferences" and then select "Notifications".

We will use **Piazza** for technical discussions between students, TAs, and lecturers.

Submitting Assignments

Homework and lab assignments will typically be due at 11:55PM on the due dates. For specific assigning dates, due dates and grace period deadlines, refer to the course Sakai website.

Homework assignments and lab reports will be submitted as pdf files through the Assignments tool in Sakai. The code of lab assignments will be submitted to our servers. The details will be given during class.

All the assignments should be submitted prior to the deadline and must be submitted by the grace period deadline (a penalty may be applied). Submissions will not be accepted after the grace period.

Deadline Extensions

Individual extensions to homework or lab assignment deadlines may be granted in justifiable situations (e.g., participation in a University-sponsored extracurricular activity, interviews, travel for religious observances) when I have been informed in a timely manner – which is typically at least one week in advance, though may vary depending on the circumstances. To request an extension, send me an email with

- the assignment for which you are requesting an extension,
- an explanation of the circumstances necessitating the request, and
- the proposed, extended, deadline.

I will respond to your email. If the extension is granted then attach a copy of our email granting the extension to the front of your assignment and submit your assignment by the extended deadline. Work will not be accepted after the extended deadline.

Extensions are not granted for "too much work."

No extension request will be accepted within 24 hours before the deadline except for emergence circumstances.

If you ask me for an extension in person, I will direct you to send me an email as indicated above.

Extracurricular Activities

Given appropriate and timely notification, I will make every effort to assist you in working around extracurricular commitments (e.g., participation in University athletics, interviews) that might otherwise conflict with course commitments other than exams and project presentations. See "Deadline Extensions" above.

Regrades

Errors, oversights, and misinterpretations may occur. If there is an error in your grade (e.g., the total number of points incorrectly summed) or you feel that the grade you received is not commensurate with your solution, then submit to the lecturer, within one week of the graded work being returned to the class:

- your assignment in question, with
- a brief written description of the error, and
- your Duke NetID.

I will respond to your regrade request by email and make arrangements to return your work to you.

As a matter of policy, when you request a regrade you are agreeing that the grading of the entire assignment may be re-evaluated.

Course Policies

Academic Integrity

All members of the Duke community, ourselves included, are expected to abide by the Duke Community Standard (DCS). https://studentaffairs.duke.edu/conduct/about-us/duke-community-standard

Duke University is a community dedicated to scholarship, leadership, and service and to the principles of honesty, fairness, respect, and accountability. Citizens of this community commit to reflect upon and uphold these principles in all academic and nonacademic endeavors, and to protect and promote a culture of integrity.

To uphold the Duke Community Standard:

- *I will not lie, cheat, or steal in my academic endeavors;*
- *I will conduct myself honorably in all my endeavors; and*
- *I will act if the Standard is compromised.*

Suspected violations of the DCS will be forwarded to the Office of Student Conduct where they will be given due process and dealt with in accordance with University policies. Common sanctions for being found responsible for academic dishonesty range from probation to a two-semester suspension, depending on the circumstances. Sanctions often include required reading and a piece of reflective writing about the experience.

You may encounter situations where whether or not an action would be considered a violation of the DCS is not entirely clear to you. A guiding principle is, if you wouldn't want anyone, or everyone, to know you did it, then don't do it.

If you are unsure whether an action would be considered a violation of the Duke Community Standard, please talk with me before you do it.

Students most often experience lapses in judgment when they find themselves under time pressure to complete assignments. You can help yourself by starting your assignments well in advance of when they are due. It is never to your advantage to take academically dishonest shortcuts, as doing so will invariably lead to an outcome worse than either submitting an incomplete assignment or not submitting it at all.

Collaboration

I recognize that cooperative learning has the potential to enhance and enrich your learning experience, and as such encourage you to engage with your fellow students in group discussions of concepts and problem-solving approaches. While group discussions of concepts and problem-solving approaches for comparison of final numerical answers are acceptable, sharing your solution as an answer to the question, "Can you help me with problem 3?" is not acceptable. Although you may discuss the general principles and methods taught in the course and how they relate to the problem at hand, you may not share your specific solution. Do not view or copy another student's solution, and do not allow another student to view or copy your solution.

Helpful tips to avoid violations of academic integrity:

- Do not share your work with other students, even if it has already been graded and returned to you – other students may have deadline extensions.
- Thoughtfully consider the nature of information exchange with other students, and be especially aware that the pressure of an impending deadline can alter your judgment.

Attendance

Attendance at lecture and laboratory is mandatory.

You are responsible for everything covered in lecture and labs; absence from lecture or lab does not excuse knowledge of material covered or grading standard.

Courtesy

Arrive before the start of class/lab and configure electronic devices so they are completely silent.

Discrepancy

If there is a discrepancy of the guidance between the TAs and the lecturer, the final explanation right belongs to the lecturer.