Fall 2023 Deep Learning: Syllabus and Schedule

Schedule	Syllabus	Lecture Slides	Piazza

Course Description:

This course is an introduction to deep learning, a branch of machine learning concerned with the development and application of modern neural networks. Deep learning algorithms extract layered high-level representations of data in a way that maximizes performance on a given task. For example, when asked to recognize faces, a deep neural network may learn to represent image pixels first with edges, followed by larger shapes, then parts of the face like eyes and ears, and, finally, individual face identities. Deep learning is behind many recent advances in AI, including Siri's and Alexa's speech recognition, Facebook's tag suggestions and self-driving cars. We will cover a range of topics from basic neural networks, convolutional and recurrent network structures, deep unsupervised and reinforcement learning, and applications to problem domains like speech recognition and computer vision.
Prerequisites: a strong mathematical background in calculus, linear algebra, and probability & statistics, as well as prior coursework in machine learning and programming experience in Python.

Lecture:

CAS CS523 A2/ ENG EC523 A1
M/W 2:30-4:15pm, PHO 203

Instructor:

Kayhan Batmanghelich
office hours: M 10-11, outside PHO 421

Teaching Assistant:

Li Sun (lisub@bu.edu),
Office hours: Thursday 2pm – 3pm,
Location PHO 4th floor

Graders/Additional Staff:

Jordan Koseski
Priyank Negi

How to Contact Us: Please use Piazza for all communication; if your question is only directed to the instructors, please make a post to "Individual Student(s) / Instructor(s)" and select "Instructors".

Piazza: https://piazza.com/bu/fall2023/ec523cs523a1
We will be using piazza for online discussions, questions, and posting assignments.

Gradescope: we will be using Gradescope for submitting and grading assignments.

Schedule*

	Topic (Instructor)	Details	Homework
Wed Sep 6	1. Course overview	What is deep learning? DL successes; syllabus & course logistics.	
Mon Sep 11	2. Math/ML Review I	Probability, distributions, maximum likelihood, empirical risk	HW1 out

		minimization.	(machine learning prereq)
Wed Sep 13	3. Math/ML Review II	Generalization, train/validation/test splits, stability, stochastic gradient descent.	
Mon Sep 18	4. Neural network basics I	Classification and regression tasks, perceptron, universal approximation.	
Wed Sep 20	5. Neural network basics II	MLP, activation functions, surrogate loss functions, softmax, and compression.	
Mon Sep 25	6. Neural network basics III	Automatic differentiation and backpropagation, matrix derivatives.	HW1 Due HW2 out
Wed Sep 27	7. Training I	Mini-batching, regularization, adversarial examples, dropout, batch norm, layer norm.	
Mon Oct 2	8. Training II	Momentum and acceleration, physical interpretation of accelerated gradient descent, stochastic gradients and variance.	SCC Info
Wed Oct 4	9. Training III	Adaptive gradient methods: adagrad, adam, Lars/Lamb, and large batch sizes.	
Mon Oct 9	NO CLASS		HW2 due HW3 out
Wed Oct 11	10. CNNs	Convolutional neural networks, including AlexNet, VGG, and Inception; **Reading**: Goodfellow Ch9.1-9.3.	
Mon Oct 16	11. CNNs II and Advanced Architecture Design	Modern Conv Nets, ResNet	Project Proposal Due Template How to make a group submission
Wed Oct 18	12. Advanced Architecture Design II		
Mon Oct 23	13. Deep Unsupervised Learning I	Autoencoders	HW3 due
Wed Oct 25	14. Deep Unsupervised Learning II	Variational Autoencoders	
Mon Oct 30	15. Deep Unsupervised Learning III	Diffusion Models	HW 4 out
Wed Nov 1	16. Deep Unsupervised Learning IV	Generative Adversarial Networks	
Mon Nov 6	17. RNNs	Recurrent neural networks; sequence modeling; backpropagation through time; vanishing/exploding gradient problem; gradient clipping, long-short term memory (LSTM).	
Wed Nov 8	18. Transformers I	Embeddings, word vectors, self-attention, transformers.	

Mon Nov 13	19. Transformers II	GPT, BERT, pretraining, masked language modeling task, few-shot learning.	HW4 due HW5 out
Wed Nov 15	20. Deep Reinforcement Learning I	Overview of RL, Policy Gradient	Project Status Report Due Thu 11:59pm Template
Mon Nov 20	21. Deep Reinforcement Learning II	Actor-Critic, Q-learning	
Wed Nov 22	NO CLASS		
Mon Nov 27	22. Explainability I		
Wed Nov 29	23. Explainability II	Automatic Speech Recognition	
Mon Dec 4	24. Self-supervised Learning	self-supervised learning (slides from this tutorial)	
Wed Dec 6	25. TBD		HW 5 due
Mon Dec 11	Project Presentations I	Presentation Schedule Send slides (or slide link) to my email	Slides due Thu Dec 11 12:00pm NOON Slide Template
Wed Dec 13	Project Presentations II	Presentation Schedule Send slides (or slide link) to my email	Slides due Thu Dec 13 12:00pm NOON Slide Template
Mon Dec 18	Final Project reports/code due (No lecture)	Upload your reports on GradeScope, **use the template** Sample Projects: reports, slides	Due Mon 11:59pm MIDNIGHT Report Template

*schedule is tentative and is subject to change.

Syllabus

Course Prerequisites

This is an upper-level undergraduate/graduate course. All students should have the following skills:

- Calculus, Linear Algebra
- Probability & Statistics
- Ability to code in Python
- Background in machine learning (e.g. EC 414, EC 503, CS 542)

Textbook

The recommended textbook for the course is

- Ian Goodfellow, Yoshua Bengio, Aaron Courville. Deep Learning. MIT Press, 2016.

This book is available online, and need not be purchased. Another recent book is

- Aston Zhang, Zack C. Lipton, Mu Li, and Alexander Smola. Dive into Deep Learning, 2020.

Other recommended supplemental textbooks on general machine learning:

- Duda, R.O., Hart, P.E., and Stork, D.G. **Pattern Classification**. Wiley-Interscience. 2nd Edition. 2001.
- Theodoridis, S. and Koutroumbas, K. **Pattern Recognition. Edition 4**. Academic Press, 2008.
- Russell, S. and Norvig, N. **Artificial Intelligence: A Modern Approach**. Prentice Hall Series in Artificial Intelligence. 2003.
- Bishop, C. M. **Neural Networks for Pattern Recognition**. Oxford University Press. 1995.
- Hastie, T., Tibshirani, R. and Friedman, J. **The Elements of Statistical Learning**. Springer. 2001.
- Koller, D. and Friedman, N. **Probabilistic Graphical Models**. MIT Press. 2009.

Recommended online courses

- http://cs231n.stanford.edu/ CS231n: Convolutional Neural Networks for Visual Recognition
- http://web.stanford.edu/class/cs224n/ CS224n: Natural Language Processing with Deep Learning
- http://rll.berkeley.edu/deeprlcourse/ CS 294: Deep Reinforcement Learning
- http://distill.pub/ Very nice explanations of some DL concepts

Deliverables/Graded Work

There will be five homework assignments, each consisting of written and/or coding problems, and a final project. Homework grade will be based on a randomly selected subset of questions (the same for everyone). The worst homework grade will be dropped. The project will be done in teams of 3-4 students and will have several deliverables including a proposal, progress update(s), final report and a final in-class/virtual presentation. The course grade consists of the following:

- Homeworks (hw1 and best 3 of 2-5) 50%
- Project (including all components) 45%
- Class/Piazza participation 5%

Software/Hardware

Programming assignments and projects will be developed in the Python programming language. We will also use the pytorch deep learning library for some homeworks and for the project. Students are expected to use the Shared Computing Cluster (SCC) and/or their own machines to complete work that does not require a GPU. For the projects, we will provide GPU resources.

If you do not already have a CS account and would like one, you should stop by the CS undergraduate lab (EMA 302) and activate one. This process takes only a few minutes and can be done at any time during the lab's operating hours: <http://www.bu.edu/cs/resources/laboratories/undergraduate-lab/>

Late Policy

Late work will incur the following penalties

- Project deliverables: 20% off per day up to 2 days
- Homework 20% off per day, up to 3 days
- We will automatically drop the lowest scoring homework (except hw1)

Academic Honesty Policy

The instructors take academic honesty very seriously. Cheating, plagiarism and other misconduct may be subject to grading penalties up to failing the course. Students enrolled in the course are responsible for familiarizing themselves with the detailed BU policy, available here. In particular, plagiarism is defined as follows and applies to all written materials and software, including material found online. Collaboration on homework is allowed, but should be acknowledged and you should always come up with your own solution rather than copying (which is defined as plagiarism):

> **Plagiarism**: Representing the work of another as one's own. Plagiarism includes but is not limited to the following: copying the answers of another student on an examination, copying or restating the work or ideas of another person or persons in any oral or written work (printed or electronic) without citing the appropriate source, and collaborating with someone else in an academic endeavor without acknowledging his or her

contribution. Plagiarism can consist of acts of commission-appropriating the words or ideas of another-or omission failing to acknowledge/document/credit the source or creator of words or ideas (see below for a detailed definition of plagiarism). It also includes colluding with someone else in an academic endeavor without acknowledging his or her contribution, using audio or video footage that comes from another source (including work done by another student) without permission and acknowledgement of that source.

Religious Observance

Students are permitted to be absent from class, including classes involving examinations, labs, excursions, and other special events, for purposes of religious observance. In-class, take-home and lab assignments, and other work shall be made up in consultation with the student's instructors. More details on BU's religious observance policy are available here.