Applied Deep Learning, Fall 2020
Syllabus and FAQ

Day / Time: Thursday, 7:00pm-9:30pm
Where: Hybrid, 501 Schermerhorn Hall
Instructor: Josh Gordon, joshua@cs.columbuia.edu
Office hours: See CourseWorks

Course description

This course provides a practical introduction to Deep Learning. We aim to help you understand the fundamentals of neural networks (DNNs, CNNs, and RNNs), and prepare you to successfully apply them in practice. This course will be taught using open-source software, including TensorFlow. In addition to covering the fundamental methods, we will discuss the rapidly developing space of frameworks and applications, including deep learning on mobile and the web, and applications in healthcare. This course emphasises fairness, responsibility, and testing, and teaches best practices with these in mind.

Syllabus

Subject to change based on your interests.

Date	Where	Topic(s)
9/10	Online only	**Intro to Deep Learning** Topics: Deep Learning overview
9/10	N/A	Assignment 1 released (due 9/29)
9/17	Online only	**AI Responsibility; linear models** Weights, activations, softmax, cross entropy. Famous data-driven incidents, and why they occurred.
9/24	Online only	**Working with images** Convolution and CNNs, transfer learning, data augmentation, famous architectures, image segmentation, style transfer, deep dream. Loading images.
9/24	N/A	Assignment 2 released (due 10/15)
10/1	Online only	**Medical imaging case studies** Recent work in depth
10/8	On-campus	**Trusting and explaining models** NLP basics; Hidden Technical Debt in ML systems; What's your ML test score.

10/15	Online only	**LIME, Integrated Gradients**, and **Review session (will cover first half of midterm)**
10/15	N/A	Assignment 3 released (due 10/29)
10/22	On-campus	**Optimization; Writing efficient input pipelines; Distributed training** Gradient descent, backpropagation, optimizers, batchnorm, dropout; caching, prefetching; Distributed training / all reduce / data and model parallelism.
10/29	N/A	Project proposal posted (due 11/5)
10/29	N/A	Assignment 4 released (due 11/12)
11/5	On-campus	**Working with sequences** RNNs, sequence classification, sequence generation.
11/12	On-campus	**Generative models; Adversarial examples** Autoencoders, GANs, fooling classifiers
11/12	N/A	Assignment 5 released (due 12/3)
11/19	On-campus	**Working with structured data; TinyML; Federated Learning; SSL** Handing structured data and imbalanced data, federated learning & privacy; Semi-supervised and self-supervised DL>
11/26	N/A	No class (University holiday)
12/3	Online only	Midterm (you will have the entire period)
12/10	Online only	**Guest lecture & Attention, Transformers, and BERT** As above
12/22	N/A	Final project due

Course assistants

All office hours with the course assistants are virtual, and will take place over Zoom.

- Heetika Vipul Gada <hg2532@columbia.edu>
- Kartik Parnami <kp2844@columbia.edu>
- Michelle Maria Roy <mr3966@columbia.edu>
- Prajwal Prakash <pp2719@columbia.edu>
- Su Ji Park <sp3581@columbia.edu>
- Jiayin Yang <jy3016@columbia.edu>
- Priyal Aggarwal <pa2564@columbia.edu>
- Mansi Khemka <mk4282@columbia.edu>

FAQ

Prerequisites and difficulty level
Data science students and practitioners come from diverse academic backgrounds (computer science, medicine, bme, chem, etc). It's important to the instructor that this course is accessible.

- You should have previously taken COMS W4721 (Machine Learning for Data Science), or an equivalent introductory machine learning course. We will refresh a few concepts as we go.

- You should be comfortable programming in Python (including NumPy and Matplotlib). The Python Data Science Handbook is a helpful reference (the content is available for free online).

This course is right for you if you're interested in developing practical skills, and learning how deep learning is used in industry. The department also offers theoretical courses if those match your interests.

Workload and grading
- A course project and presentation (30%)
- 5 practical homework assignments (50%)
- A midterm (20%), mostly focused on topics like Responsible AI, fairness, and applications.

Course project
The course project gives you an opportunity to create a portfolio of work. Students may choose to complete a suggested project, or propose their own. The suggested project involves developing a solution to the CAMELYON16 challenge; starter code and data will be provided. Note that it's relatively easy to develop a basic solution, but difficult to develop a good one.

The project may be done in groups of up to two. If you're doing a group project, the grading expectations will be higher. Deliverables include:
- A link to your GitHub repo.
- A YouTube video (+/- 15 minutes, including a recorded presentation, demo, and code walkthrough). Your video can be unlisted, and does not need to be shared publicly. If you are working in a group, each student must present equally.

Textbooks
There are two. Students should purchase a copy of Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow, 2nd Edition[1]. It's full of practical advice. We will supplement this book with Deep Learning by Ian Goodfellow (available for free online).

[1] Be sure you have the 2nd edition (the 1st edition uses an earlier version of TensorFlow).

Software and languages
There are many open-source deep learning frameworks available today. We will use TensorFlow 2.0. After learning the fundamentals in TensorFlow, you should be able to move to any major framework that exists today, including PyTorch and JAX. All assignments will be in Python, with a small amount of TensorFlow.js in one (prior JavaScript experience is not assumed). You will also have the option to use TensorFlow Lite and/or TensorFlow Lite Micro in your final project.

Programming environment
We recommend Colaboratory (a web-based Jupyter environment that includes a free GPU). This will be sufficient for the homework assignments in this course. Alternatively, you can install everything locally, use the cloud platform of your choice, or sign up for Colab Pro (this costs $10 / month).

Late policy
The penalty for late assignments is 10% / day.

Class format
We have a long time slot. We will use some class time for demos, discussion, to start on the homework, and for additional office hours.

Collaboration policy
Feel free to study in groups. You may discuss your approach to homework assignments, and help each other debug. That said everyone must write and submit their own code (with the exception of the course project, if you are working in a group). Please keep the university's academic integrity policy in mind.

Exam policy
The midterm is closed-book and closed-notes. If you are traveling for work, and/or have an interview scheduled, it's fine to take it early - please send the course assistants an email and they will help you schedule a time.

Disability Services
Disability Services facilitates equal access for students with disabilities by coordinating accommodations and support services, cultivating a campus culture that is sensitive and responsive to the needs of students. Students seeking accommodations or support services from Disability Services are required to register with the office. If you are interested in pursuing an evaluation for a learning disability, please visit the referrals and other campus resources page.