Course Overview and Goals

A Deep Learning model

- is a combination of components
- each component parameterized
- is fit (finding optimal values for parameters) based on a training dataset
- by optimizing a Loss Function

In the Intro course

- things were "small": datasets, number of parameters
- components were assembled in a simple but rigid order: sequence of Layers (the "Sequential" architecture)
- Loss functions were few and pre-specified: Mean Squared Error, Cross Entropy

In the Advanced course we depart from the simplifications of the Intro course

- How to use Big Data in Small Memory
- Deal with models with parameters numbering in the billions
- Introduce an unrestricted (technically: Directed Acyclic Graph) architecture for organizing components (the "Functional" model)
- Re-using advanced models that are too big to construct ourselves
- Loss functions designed to express the semantics of the task to be solved

The course combines two approaches

- Establishing the technical prerequisites for understanding complex, advanced Deep Learning models
  - Preparing the student for a Final Project involving coding using an advanced model
- A survey of interesting (often cutting-edge) topics in Deep Learning
  - Emphasis is on breadth rather than depth: many topics covered at a very high level
    - Goal is to enable the student to learn about what is at the leading-edge of Deep Learning so that motivated students can be guided in further self-study

Course requirements

- Graduate standing
- FRE 7773 Intro to Machine Learning
- Solid background in Deep Learning. Can be satisfied in a number of ways
- The Deep Learning lectures of the Instructor’s section of FRE 7773 will provide this background
- OR Equivalent knowledge
  - consult the course materials (available on Github) for the Instructor’s section of FRE 7773

### Course materials

#### Lectures
Lectures will be distributed in advance of each class as a Jupyter notebook.

On the class Brightspace page, there is a section entitled “Week 0” with preparation to be performed prior to the first class.

1. Setting up your Machine Learning environment, e.g. Jupyter
2. Obtaining lesson materials

#### Textbooks
Most of the reading material will be research papers/publications available on-line. These will be presented as links in the weekly Jupyter notebooks.

The following books are useful references but will not be used directly in the course:

2. Deep Learning, by Ian Goodfellow, Yoshua Bengio, and Aaron Courville
   1. Also available as a free, online book [https://www.deeplearningbook.org/](https://www.deeplearningbook.org/)

#### Grading
95% of the final grade will be based on the Course Project, due toward the end of the course. The remaining 5% will be based on class participation, which is highly encouraged.

This material is very new; we are all on a shared journey and can benefit from mutual insights.

#### Disability Disclosure Statement
Academic accommodations are available for students with disabilities. Please contact the Moses Center for Students with Disabilities (212-998-4980 or mosescsd@nyu.edu) for further information.
Students who are requesting academic accommodations are advised to reach out to the Moses Center as early as possible in the semester for assistance.

**Inclusion statement**

The NYU Tandon School values an inclusive and equitable environment for all our students. I hope to foster a sense of community in this class and consider it a place where individuals of all backgrounds, beliefs, ethnicities, national origins, gender identities, sexual orientations, religious and political affiliations, and abilities will be treated with respect. It is my intent that all students’ learning needs be addressed both in and out of class, and that the diversity that students bring to this class be viewed as a resource, strength and benefit. If this standard is not being upheld, please feel free to speak with me.