# STAT 991: Topics in Deep Learning

Department of Statistics, The Wharton School, University of Pennsylvania, Fall 2018

---

**Time:** Thursdays 12:00pm-1:20pm **Place:** Huntsman Hall F94

---

**Objectives:** This advanced seminar course will explore several topics in deep learning. We will discuss both theory and applications.

Students will work in groups of 1-3 people and present a topic over 1-2 lectures. The presentations will summarize the work of several research papers on a topic. They will include necessary background, algorithms, as well as results and proofs (in case of theory). Finally, the goal of the course will also be to identity new research directions.

I am hoping that this course will provide a venue for discussion for interested students at Penn. I am also hoping that the most interested participants can continue this, informally, as a reading group, after the semester ends.

**Prerequisites:** You are expected to have some basic familiarity with deep learning. In particular, you should know what the following terms mean: deep net, hidden layer, activation, weight, bias, ReLU, backpropagation, SGD, dropout, batchnorm, CNN, RNN, Keras. In addition, you are expected to have a basic familiarity with statistics and machine learning.

Some parts of this will be covered in the first few lectures, and provided in the lecture notes. For the remainder, you are expected to learn it yourself. Good resources include:

1. Andrew Ng's course, https://www.coursera.org/learn/neural-networks-deep-learning. (lecture videos)

**Topics:** A list of possible topics are included below. We can cover other topics based on student interest.

- Introduction
- Common architectures, important applications: What is VGG-19? How does AlphaGo Zero work?
- Training: How can we train a deep net, if the loss is so non-convex? (Computational complexity, linear algebra, statistical physiscs)
- Generalization: Why does a net generalize, if the number of parameters is so much larger than the training data size? (Statistical learning theory)
- Architecture: Why deep convolutional networks, and why not something else? (Approximation theory, harmonic analysis)
- Uncertainty quantification: How can we evaluate the certainty in our predictions, if the model is so complex? e.g., Are you sure that the car is at least 10 feet away? (Bayesian statistics)

- Generative adversarial networks (GANs): How can we learn to generate from a complex distribution? (Nonparametric density estimation, Optimal transport)
- Interpretation: How can we understand and interpret deep nets? e.g., If a net predicts that you are a criminal, why does it think that?
- Artificial intelligence: Are deep nets intelligent? How close are we to creating AI?

A few papers will be provided as a starting point for each area. To find more papers, you can use the following resources: Google Scholar (keyword search, cited by, related papers). Curated paper lists.

**Computation:** There is no formal computational component of the course. However, it can be extremely helpful to build experience doing deep learning yourself. For this, if you are coming from a statistics background, the easiest route is to use Keras in R. If you have experience using other computational frameworks, you may use those in your presentation.

We will do some small deep learning experiments in class. For those lectures, I encourage you to bring laptops to class, so that we can run the code together.

1. To prepare, you will need to install the Python distribution Anaconda on your laptop.
2. You will also need to install Keras. An R script to do this will be posted on Canvas.
3. We will use material from the book Deep Learning in R, by Chollet and Allaire. See https://github.com/jjallaire/deep-learning-with-r-notebooks.
4. Google has generously provided an educational grant for students to use the Google Cloud Platform. If there is interest, we can try to explore this.

**Instructor:**

- Edgar Dobriban, dobriban@wharton.upenn.edu, Office: 465 JMHH
  Office Hours: by appointment

**Course Page:**

- Canvas, for announcements and materials: https://canvas.upenn.edu/courses/1419899
- Piazza, for discussion. https://www.piazza.com/class/jl4jlcg887u5t

**Feedback:** I am interested to hear about your experience and suggestions for the class.

**Grading Policy:** The course grade will be driven by two factors: presentation (80%), and class participation (20%). The components of each are

- Presentation: Clarity (ability of others to follow). Correctness. Coverage (did you cover the important parts). Insight.

  Think of the presentation as a course project. You will need to prepare the presentation (slides to be presented in class, or cca 6 pages of lecture notes) and provide it to the class 24 hours in advance. These will be posted on Canvas.

We are aiming that the first student presentations be on September 13. There is a Google Sheet where you can sign up for specific topics and dates.

I can guide with choosing topics, literature search, and structuring the presentation.

- Class participation: Attendance. Reading the presentation materials ahead of class, and writing a one-paragraph summary.