

# STAT GR 5242, Advanced Machine Learning

## Schedule:

**Time:** Tue-Thu 10:10-11:25 am **Location:** Zoom.

---

### Instructors:

- Samory Kpotufe. *email:* skk2175@columbia.edu, *Office hours:* 30 mn right after each class.
- Kamiar Rahnama Rad. *email:* kamiar.rad@baruch.cuny.edu, *Office hours:* TBD

### Assistant Instructors:

- Gan Yuan. *email:* cixiyg@gmail.com, *Office hours:* 9-10:30pm Thursday
  - Zhen Huang. *email:* zh2395@columbia.edu, *Office hours:* 1:30-3pm Monday
- 
- 

## Evaluation:

Homeworks and projects, some of which allow collaboration. The assignments in each half of the course will count for 50% of the grades.

---

---

## Course Overview:

The course is made up of 2 parts, 6 weeks each. The first part, more conceptual, covers various challenges encountered in the practice of ML, and aims at yielding a strong sense of solutions and principles used in addressing them. The second part covers a wide range of deep-learning approaches to common ML problems, and aims to equip students with an advanced toolkit for real-world ML applications.

**Companion Course** (GR5245, Instructor Ka Yi Ng, covering Python and TensorFlow):

<http://www.math.columbia.edu/~kyn/GR5245%20syllabus%20Fall%202021.pdf>

The next pages give a more detailed overview of subjects, prerequisites and postrequisites.

## **Advanced ML (GR5242)**

### **Prerequisites:**

- Enough programming knowledge to pick up machine learning libraries like sickit-learn and tensorflow
- Basic Unsupervised Learning
  - k-means
  - PCA
- Comfortable with basic optimization
  - EM
  - gradient descent
- Supervised Learning
  - Multiclass logistic regression (and dependent topics)
  - SVMs
- The concept of a latent representation
  - PCA
- Exposure to models that address dependent structure in the data
  - time series
  - HMM/markov chains

### **Post-requisites (Learning outcomes):**

- Modern approaches to hyperparameter choices both unsupervised and supervised.
- Some understanding of complexity and generalization
- Modern solutions to computational tradeoffs: SGD and Automatic differentiation
- Deep learning software: Tensorflow
- Convolutional neural networks: modern computer vision
- Recurrent neural networks: up to stacked LSTMs and modern NLP applications
- Multi-arm bandits, up to Thompson sampling
- Reinforcement Learning in the modern (deep learning) world

**Course:** Advanced Machine Learning (GR5242)

List of Topics	Importance	Time Spent
Hyperparameter Choices in Unsupervised Learning		
- Density Estimation and parameter choices (L2, Elbow, Links to GANs)	5	2 lectures
- Density based Clustering and Number of Clusters (Mean Shift, Kernel k-Means/PCA, Spectral Clustering)	4	2 lectures
Hyperparameter Choices in Supervised Learning		
- Notions of complexity and Complexity Regularization	5	2 lectures
- Dimension reduction, early Stopping, Learning Rates, as complexity Regularization	4	1 lectures
- Bayesian Approaches: complexity under Priors.	4	2 lectures
Tradeoffs in Constrained Settings		
- Subsampling and Data Quantization (hyperparameter choices, variance correction)	5	1 lecture
- Nystrom methods as advanced subsampling/dimensionality reduction (hyperparameter choices again are harder)	5	2 lectures
(part 2)		
Deep Learning basics (2 weeks)		
- Tensorflow, fully connected networks	5	2 lectures
- autodiff	4	1 lecture
- stochastic gradient descent	4	1 lecture
Convolutional neural networks		
- fundamentals	5	1 lecture

- using tf and cnn	5	1 lecture
- from basics to imagenet (lenet, inception, resnet)	5	1 lecture
- transfer learning	5	1 lecture
Reinforcement learning		
- Basics, multi-arm bandits	5	1 lecture
- Dynamic programming, td learning, q learning	5	1 lecture
- Deep q learning, modern deep rl	4	1 lecture
Natural language processing		
- From bigrams to RNNS	5	1 lecture
- From RNNs to LSTM/GRUs	5	1 lecture
- Modern NLP: stacked LSTM, embeddings, etc	4	1 lecture
Representation learning		
- VAE and GAN	4	1 lecture