

## STAT 344NE: Introduction to Neural Networks

**About the Course****Instructor** Evan RayEmail: [eray@mholyoke.edu](mailto:eray@mholyoke.edu)

Office: Clapp 404C

Office Hours: I will hold regularly scheduled office hours each week at times to be selected by you. These times will be posted on the course web site. Please do not hesitate to contact me to set up appointments for additional office hours at any time!

**Classes**

Mon, Wed, Fri 11:00am - 12:15 pm in Clapp 420.

**Please plan to bring your laptop to class.**

**Course Website** The course website is at [http://www.evanlray.com/stat344ne\\_s2020/](http://www.evanlray.com/stat344ne_s2020/). I will update it regularly with lecture notes and materials used in class. Lab assignments we will work through in class and homework assignments will be distributed on GitHub.

**Description** Neural networks are among the most commonly used models in industry, and are particularly common in settings with large data sets or with unstructured data such as image data. In this course, we will develop neural networks as statistical models that generalize linear regression and logistic regression. We will emphasize (a) the statistical formulation of models and estimation by maximum likelihood; (b) a mathematical development of underlying ideas and methods such as gradient descent and regularization; and (c) computational skills necessarily to effectively implement and use neural network models using NumPy, Pandas, and Keras.

A tentative schedule and topic list is below. In setting this schedule, my goal has been to build up an understanding of the foundations of the most commonly used methods as rapidly as possible, so that you can begin working on projects. In the later part of the semester we will go back and fill in some missing details and see more advanced ideas and applications. This schedule is subject to change; an up-to-date list of topics covered so far and a rough time line for upcoming classes will be kept on the course website.

**Textbooks** We will use two text books for this course:

1. “Deep Learning with Python” by François Chollet. This book is a great practical guide to working with neural networks in Keras, but it has limited mathematical depth. A copy is on reserve at the library and an ebook version is linked to from the course website.

2. “Deep Learning” by Ian Goodfellow, Yoshua Bengio, and Aaron Courville. This book has limited information about practical implementation of neural networks, but it covers the mathematics in depth.

Unit	Topic	Weeks
Foundations	<b>Structure:</b> Structure of neural network models; log-likelihood loss functions for regression, binary classification, and multi-class classification tasks; common activation functions	1, 2
	<b>Estimation:</b> Gradient descent and stochastic gradient descent; back-propagation; $L_2$ regularization; train/validation/test splits and hyperparameter tuning	2, 3
	<b>Computation:</b> Defining and fitting simple models with NumPy and Keras; managing data with Pandas	1 – 4
Network Architectures	<b>CNNs:</b> convolutions and pooling; common architectures including AlexNet, VGG, GoogLeNet, ResNet; transfer learning	5, 6
	<b>RNNs:</b> RNN; GRU; LSTM	7, 8
	<b>Computation:</b> Implementing basic versions of CNNs and RNNs with NumPy; implementing common models with Keras	5 – 8
More Detail on Estimation:	<b>Optimization:</b> gradient descent with momentum; RMSprop; Adam	10, 11
	<b>Other Tricks:</b> early stopping; dropout; batch normalization	11
Special Topics	<b>Case Studies:</b> Examples to be selected from: object detection with YOLO; neural style transfer; variational autoencoders; generative adversarial networks; machine translation; image captioning; etc.	12 – 14

**Time commitment** While the exact time commitment for the class will vary individually and over the course of the semester, I recommend that you budget approximately three out-of-class hours for every class hour to complete the reading, homework, and project. It should be feasible to satisfactorily complete the requirements with approximately twelve hours per week of time commitment. If you are spending more time than this on a regular basis I would encourage you to check in with me.

## Policies

**Attendance** Your attendance in class is crucial, unless you are sick. If you are sick, please let me know and stay home and rest; I hope you feel better!

**Collaboration** Much of this course will operate on a collaborative basis, and you are expected and encouraged to work together with a partner or in small groups to study, complete homework assignments, and prepare for exams. However, every word that you write must be your own. Copying and pasting sentences, paragraphs, or large blocks of code from another student is not acceptable and will receive no credit or a penalty. No interaction with anyone but the instructor is allowed on any exams or quizzes. All students, staff and faculty are bound by the Mount Holyoke College Honor Code.

To sum up: On homeworks and labs, **I want you to work together.** *But, you must write up your answers yourself.*

Cases of dishonesty, plagiarism, etc., will be reported.

## Assignments

Your grade for this course will be a weighted average of scores from several components:

Item	Weight
Participation and Labs	5%
Homework	15%
Quizzes	15%
Midterms	40%
Final Project	25%

**Participation and Labs** The best way to learn statistics is to do it. This class will include some interactive element, whether it be a worked example or a computer lab, nearly every day. Although I will not grade these for correctness, I expect you to complete them. I will occasionally look at submitted labs to see how everyone is doing and whether there are any points I need to address in class. I am always happy to answer any questions you have about these in-class exercises and labs.

**Homework** We will have regular homework assignments to be completed outside of class. Occasionally, questions may relate to material in the reading that will not be covered in class.

**Exams** There will be two or three in-class “midterm” exams and occasional in-class quizzes. The first midterm will cover material from the “Foundations” unit; the second will cover material from the “Network Architectures” and “More Detail on Estimation” units; and a possible third midterm will cover material from the “Special Topics” unit. We will not have a cumulative final, but we may have a late-in-the-semester midterm. Midterms will definitely have an in-class written component, and may also have a take home component. There will be a quiz most weeks that we don’t have a midterm. Quizzes will always be announced at least once class session in advance, and midterms at least one week in advance.

No communication with anyone besides the instructor is allowed on these assessments.

**Project** A large component of the course will be a project which will be presented to your classmates. Briefly, this project will entail application of a neural network model to a problem of your choice. A separate handout will provide additional details.

**Grading** When grading your written work, I am looking for solutions that are technically correct and reasoning that is clearly explained. *Numerically correct answers, alone, are not sufficient* on homework, tests or quizzes. Neatness and organization are valued, with brief, clear answers that explain your thinking. If I cannot read or follow your work, I cannot give you full credit for it.