**Machine Learning 410**

**Deep Learning**

**March, 2018**

**Course Objectives**

This course will provide you with a background in an exciting relatively new area of machine learning and artificial intelligence know as deep learning. The focus is hands-on applications of deep learning models for unstructured data. You will achieve the following in this course:

1. Develop and apply an understanding of the theory of deep learning. Specific areas in which you will develop understanding include:
   1. Algorithms for training deep learning models.
   2. Representations used in deep learning models.
2. Apply the deep learning models to large scale and complex unstructured data problems, with a focus on:
   1. Image understanding.
   2. Natural language processing (NLP)

**Outline**

This course will follow generally follow the outline below. The exact topics to be covered will be adjusted as the course progresses.

1. **Lesson 1: Introduction to Deep Learning - Steve**
   1. What is deep learning and why do we care?
   2. Why is representation so important in AI?
   3. Introduction to feed forward networks.
   4. Introduction to loss functions
   5. Activation functions
   6. Introduction to back propagation and the chain rule
   7. Introduction to Keras and working with tensors
   8. Homework/lab: Regression with a deep feed forward network
2. **Lesson 2: Regularization for Deep Learning - Steve**
   1. Overfitting and Regularization
   2. L1 and L2 regularization
   3. Early stopping
   4. Dropout regularization
   5. Loss functions and information theory
   6. Measuring convergence with Keras
   7. Deep network architecture, depth vs. width
   8. Homework/lab: Sentiment analysis with IMBD reviews
3. **Lesson 3: Optimization for Deep Learning - Steve**
   1. Gradients and optimization
   2. Second order methods
   3. Adaptive optimization methods
   4. Scaling and stochastic gradient decent
   5. RMSProp
   6. Homework/lab: TBD
4. **Lesson 4: Convolutional Neural Nets and Images Part 1 - Steve**
   1. Representation in neural nets
   2. Introduction to convolution operators
   3. Pooling
   4. Semi-supervised learning
   5. Working with images in Keras
   6. Homework/lab: MNIST classification
5. **Lesson 5: Recursive Neural Nets and Time Series -Steve**
   1. Introduction to recursive neural nets
   2. Recurrent NN architectures
   3. Working with time series data
   4. Text analysis with recursive neural nets
   5. Dropout regularization for recursive neural nets
   6. Homework/Lab: Predicting bicycle demand
6. **Lesson 6 – Introduction to Tensor Flow - Vadim**
   1. Overview
   2. Graphs in Tensor Flow
   3. Working with images in Tensor Flow
   4. Working with sequences in Tensor Flow
   5. Representations in Tensor Flow
7. **Lesson 7: Convolutional neural nets and images Part 2 - Vadim**
   1. Convolutional architectures
   2. Using pre-trained convolutional networks
   3. Dropout for convolutional networks
   4. Visualization of convolutional networks
   5. Deep convolutional architectures
   6. Homework/Lab: Cat and dog identification with regularization
8. **Lesson 8: Deep text analytics - Vadim**
   1. Bi-directional recursive neural nets for text
   2. Embedding models for text
   3. Convolutional neural nets for text
   4. Homework/Lab: Classification of news articles
9. **Lesson 9: Practical considerations and Advanced Architectures - Vadim**
   1. Using Tensor Board to visualize deep neural networks
   2. Avoiding vanishing gradients
   3. Renormalization in deep networks – Batch normalization
   4. Searching hyperparameter space
   5. Auto encoders (optional)
   6. Keras API and non-sequential models
   7. HomeworkLab:
10. **Lesson 10: Auto encoders - Vadim**
    1. Probabilistic PCA
    2. ICA
    3. Coding
    4. Undercomplete autoencoders
    5. Regularizing autoencoders
    6. Representation and manifolds with autoencoders
    7. Denoising and stochastic autoencoders – decoders

**Grading:**

Grading for this course is based on scores in both hands-on homework exercises (2/3 of total grade) and a final independent project (1/3 of total).

**Readings**

Required readings will be assigned from *Deep Learning* by Goodfellow, Benjgio and Courville (GCB), MIT Press, 2016. You can find free access to the chapters of GCB in html form at: <http://www.deeplearningbook.org/>

A useful supplementary book is Deep Learning with Python by Chollet, Manning, 2017. You apparently read free chapters of this book in html format at: <https://www.manning.com/books/deep-learning-with-python>

**Deep Learning Frameworks**

In this course we will use primarily use the Keras framework for deep learning models. You can find the documentation for Keras at: <https://keras.io/>

We will also use selected components from the Tensor Flow framework. Tensor Flow is a good deal more complex than Keras. A Tensor Flow Programmer’s Guide is at: <https://www.tensorflow.org/programmers_guide/>

**Github Repository**

The code and lecture material for this course can be found in Github at <https://github.com/StephenElston/MachineLearning410>

**Help with Python**

We will use Python 3 in this course. If you are not familiar with the Python language, you should plan on spending some extra time

* + Python tutorial, <https://www.tutorialspoint.com/python/>
  + Python tutorial <https://docs.python.org/3/tutorial/>
  + Python for Data Analysis, 2E, Wes McKinney, <http://shop.oreilly.com/product/0636920050896.do>
  + Tutorial on scikit-learn: <http://scikit-learn.org/stable/tutorial/index.html>