Syllabus

Instructor: Farid Alizadeh

for 26:198:685 Neural Networks and Deep Learning

Last updated on 1/21/20 at 05:46 PM

Recommended books:

1. I. Goodfellow, Y. Bengio and A. Courville, *Deep Learning*, MIT Press 2016. (Available online at http://www.deeplearningbook.org)

2. F. Chollet, Deep *Learning with Pytho*n, Manning, 2018.

The first book is for theoretical understanding, the second is to learn using the Keras software along with Tensorflow to build deep learning models.

Course Web site:

https://sakai.rutgers.edu

Student work:

Each student will have to do the following tasks:

- There will be three or four projects involving Python programming (or R if you want), using Keras and/or Tensorflow (40% of grade)
- Students must propose and implement a project to present at the end of the semester. The projects may be done individually or in teams of up to three students. Students are expected to write up a summary of the project and present them in class in about 20-30 minutes on the last of the semester (or the week after). (50% of grade)
- Class participation (10% of grade)

Course Overview:

In this course we introduce the notion of (deep) neural networks (DNN). Both theoretical and conceptual aspects, and practical notions are introduced. We start with a review of linear and logistic regression and recall some of the ideas that will continue to be needed in DNN. We will review the basic algorithms, specifically optimization techniques, used for machine learning in general, and DNN's in particular. We then delve into neural networks and cover feed-forward, convolutional and recurrent networks. We will also cover autoencoders and feature selection techniques. If there is time we also introduce graphical probabilistic models and their applications in inference and in generative models. Theoretical notions are accompanied by several detailed projects. We will use primarily the Keras neural network software along with Tensorflow. Python will be our primary programming language. We will use public repositories of data such as UC-Irvine archive, Kaggle and Amazon'e AWS for both projects and class examples.

List of Topics:

1. An overview of neural networks as a machine learning tools
2. A review of machine learning basics: Bayes decision rule, supervised learning and regression and classification, forms of data, loss functions, risk, empirical risk. Review of the maximum likelihood paradigm (DL ch. 5)
3. A review of optimization techniques: Notion of gradient, subgradient and Hessian, convex and nonconvex optimization, the gradient descent method, the stochastic gradient method, Nesterov's acceleration, conjugate gradient method, Newton and Quasi Newton methods BFGS and L-BFGS methods (DL ch. 4 and supplementary notes, Newton and quasi Newton methods will be covered if there is time)
4. Feed-forward Neural networks: The notion of layers, activation functions: estimation through function composition, back propagation algorithm for computing gradients, coping with local optima, The significance and application of regularization, adjusting the learning rate, early stopping strategy, realization with Tesnorflow and Keras, (DL ch. 6)
5. Convolutional networks: The notion of convolution, achieving approximate invariance, feature extraction, applications in pattern recognition, computer vision, natural language processing (DL ch. 9)
6. Recurrent NN and time series models and applications (DL ch. 10)
7. Autoencoders and their applications (DL ch. 14)
8. Probabilistic Models Chapter 16 & 19 & 20 (if there is enough time)