Deep Learning

IE 398/CS 398
Spring 2019

Professor: Justin Sirignano
Teaching Assistants: Logan Courtney
2018 Website: https://courses.engr.illinois.edu/ie534/fa2018/

Overview:

This course is an introduction to deep learning. Topics include convolution neural networks, recurrent neural networks, and deep reinforcement learning. The course will use PyTorch to train models on GPUs.

Deep learning is computationally intensive. This course is supported by a computational grant for 25,000 GPU node hours. This provides a unique opportunity for students to develop sophisticated deep learning models.

Grading:

35% Homeworks
35% Midterm
30% Final Project

The lowest homework grade is dropped. Late homeworks are not accepted.

Prerequisites:

Python. Basic statistics, probability, and optimization. Basic knowledge of Bash/Linux is recommended.

TensorFlow, PyTorch, and Linux/Bash:

Lectures and tutorials will cover PyTorch, TensorFlow, and Linux/Bash. Example code will be provided to students. The OpenAI Gym environment for deep reinforcement learning will also be reviewed.

Topics:

- Fully-connected and feedforward networks
- Convolution networks
- Backpropagation
• Stochastic Gradient Descent
• Hyperparameter selection and parameter initialization
• Optimization algorithms (RMSprop, ADAM, momentum, etc.)
• Second-order optimization (e.g., Hessian-free optimization)
• TensorFlow, PyTorch, automatic differentiation, static versus dynamic graphs, define-by-run
• Regularization (L2 penalty, dropout, ensembles, data augmentation techniques)
• Batch normalization
• Residual neural networks
• Recurrent neural networks (LSTM and GRU networks)
• Video recognition (two-stream convolution network, 3D convolution networks, convolution networks combined with LSTM, optical flow)
• Generative Adversarial Networks
• Deep reinforcement learning (Q-learning, actor-critic, policy gradient, experience replay, double Q-learning, deep bootstrap networks, generalized advantage estimation, dueling network, continuous control, Atari games, AlphaGo)
• Distributed training of deep learning models (e.g., asynchronous stochastic gradient descent)
• Theory of deep learning (universal approximation theorem, convergence rate, and recent mathematical results)
• Convergence analysis of stochastic gradient descent, policy gradient, tabular Q-learning

Homeworks:

See course website for a list of the homeworks.

Reading:

A list of journal and conference papers will be provided to the class.

Office Hours:

Tuesday 9:30 AM -10:30 at Transportation 210A (JS)
Thursday TBA (LC)

Final Projects:

The list of final projects will be announced soon.

Project teams should be 2-3 students.