Deep Learning

CS69000-DPL
Purdue University

Aug 20, 2018

Bruno Ribeiro
Introduction

• Course overview

• What is Deep Learning?
Course overview
Logistics

• Time and location: Mon Wed 4:30pm-5:15pm, LWSN 1106

• Instructor: Bruno Ribeiro
  ribeiro@cs.purdue.edu, LWSN 2142C,
  office hours: TBD

• Teaching assistants:
  Jianfei Gao. Office hours: HASS G50, ?

• Webpage: https://www.cs.purdue.edu/~ribeirob/courses/Fall2018

• All communications via Piazza (see course webpage)

• Prerequisites: CS57800 or CS57300 or Graduate AI
Homework (50%) Workload

• Assignments include written/math exercises, programming assignments in **Python 3, Numpy & Pytorch**

• **Datasets** given in assignments are real data (dirty), that is, missing data, incomplete data, wrong data
  
  • **Cleaning** the data is part of the learning experience

• 5 homeworks, lowest grade removed from HW average
  
  • ABSOLUTELY NO extensions
  
  • ABSOLUTELY NO late homeworks

• Homework usually due 6:00am
  
  • Only TWO PDF uploads on Blackboard
Competitions

- Sometimes you will compete with each other
- Grades: Curve of best prediction ranking

Brought to you by

![Kaggle Leaderboard](image)

![The Hunger Games](image)
Midterm & Final Project

• Midterm (25%)

• Final Project (25%):
  • Projects are individual (no groups)
  • If you have your own project (research), you must
  • There will be project presentations at the end of the semester
    • Presentation time: 5 minutes per project
    • + 5 minutes for questions
Textbook

- (Required) Goodfellow, Bengio, Courville, **Deep Learning**, MIT Press
  [https://www.deeplearningbook.org/](https://www.deeplearningbook.org/)

- The texts below are recommended but not required. Reading materials will be distributed as necessary. Reading assignments will be posted on the schedule, please check regularly.

  - Trevor Hastie, Robert Tibshirani, Jerome Friedman, **The Elements of Statistical Learning**.
  
Course goals

• **Learning objectives**

  • This is an in-depth research-oriented course on deep learning. Please consider other offers on campus if you are only interested in deep learning as a tool. A lot of our coding will be in numpy (or in pytorch as a numpy for GPU) rather than existing deep learning packages. Upon completing the course, students should be able to: Understand the mathematical foundations of deep learning, including:

    • Backpropagation and Feedforward Networks
    • Robbins-Monro-type stochastic optimization
    • Auto-encoders, Representation Learning, Vector Spaces and Manifolds
    • Neural network models for structured data and the role of canonical representations in deep learning
    • Conditional Random Fields, Neural CRFs, Neural Turing Machines
    • Markov chains and Sequence learning, Backpropagation Through Time
    • Generative models, Partition Functions, Noise-Contrastive Estimation, Variational Auto-encoders, RBMs, DBMs, GANs
    • Posterior estimation in Bayesian neural networks, Hamiltonian Monte Carlo and variants

  • Understand how to evaluate the performance of neural networks, as well as formulate and test hypotheses
  • Understand how algorithmic elements interact to impact performance
Academic Honesty Policy

• **NO PART OF THE STUDENT'S ASSIGNMENT SHOULD BE COPIED FROM ANOTHER PERSON OR STUDENT (Plagiarism).** We encourage you to interact amongst yourselves: you may discuss and obtain help with basic concepts covered in lectures or the textbook, homework specification (but not solution), and program implementation (but not design). However, unless otherwise noted, work turned in should reflect your own efforts and knowledge. Sharing or copying solutions is unacceptable and could result in failure. We use copy detection software, so do not copy code and make changes (either from the Web or from other students). You are expected to take reasonable precautions to prevent others from using your work.

• **ALL MATERIALS THAT SIGNIFICANTLY CONTRIBUTED TO YOUR SOLUTION MUST BE CITED.** Say, “This part of the code was copied from Stackoverflow from page …:” Failure to do so will result in an automatic F.

• Any student not following these guidelines are subject to an automatic F as final grade. I am VERY serious about enforcing academic honesty.
Academic Honesty: First Question of Your Homework

Q0 (0pts correct answer, -1,000pts incorrect answer: (0,-1,000) pts): A correct answer to the following questions is worth 0pts. An incorrect answer is worth -1,000pts, which carries over to other homeworks and exams, and can result in an F grade in the course.

(1) Student interaction with other students / individuals:
   (a) I have copied part of my homework from another student or another person (plagiarism).
   (b) Yes, I discussed the homework with another person but came up with my own answers. Their name(s) is (are) ____________________________
   (c) No, I did not discuss the homework with anyone

(2) On using online resources:
   (a) I have copied one of my answers directly from a website (plagiarism).
   (b) I have used online resources to help me answer this question, but I came up with my own answers (you are allowed to use online resources as long as the answer is your own). Here is a list of the websites I have used in this homework:
      _________________________________________________________________
   (c) I have not used any online resources except the ones provided in the course website.
What is Deep Learning?
Deep Learning

• Different views and definitions

• My view:
  • Learn data representations with deep graphical models
    • Generally with the help of differentiable programming libraries
    • Models tend to be as general as possible, much less hand-design than traditional graphical models
  • Learning challenges due to model complexity