CSCI 5922, Spring 2020

Neural Networks and Deep Learning

Monday, Wednesday 4:30 PM - 5:45 PM

HUMN 1B80

Instructors

Dr. Adam Bloniarz, Software Engineer, Google
`adam.bloniarz@colorado.edu`

Dr. Shumin Wu, Software Engineer, Google
`shumin.wu@colorado.edu`

Teaching Assistants

Mohd Aaquib Tabrez
`mohd.tabrez@colorado.edu`

Camden Elliott Williams
`camden.elliottwilliams@colorado.edu`

Office hours

Camden: Fridays 2 - 2:55 PM in ECAE 190

Aaquib: Thursdays 11 AM - 1 PM in ECAE 172

Instructors: Wednesdays 3-4 PM in ECOT 832

CSCI 5922, Spring 2020

Neural networks have enjoyed several waves of popularity over the past half century. Each time they become popular, they promise to provide a general purpose artificial intelligence–a computer that can learn to do any task that you could program it to do. The first wave of popularity, in the late 1950s, was crushed by theoreticians who proved serious limitations to the techniques of the time. These limitations were overcome by advances that allowed neural networks to discover internal representations, leading to another wave of enthusiasm in the late 1980s. The second wave died out as more elegant, mathematically principled algorithms were developed (e.g., support-vector machines, Bayesian models). Around 2010, neural nets had a third resurgence. What happened over the past 20 years? Basically, computers got much faster and data sets got much larger, and the algorithms from the 1980s—with a few critical tweaks and improvements—appear to once again be state of the art, consistently winning competitions in computer vision, speech recognition, and natural language processing. The many accomplishments of the field have helped move research from academic journals into systems that improve our daily lives: apps that identify our friends in photos, automated vision systems that match or outperform humans in large-scale object recognition, phones and home appliances that recognize continuous, natural speech, self-driving cars, and software that translates from any language to any other language.

In this course, we'll examine the history of neural networks and state-of-the-art approaches to deep learning. Students will learn to design neural network architectures and training procedures via hands-on assignments. Students will read current research articles to appreciate state-of-the-art approaches as well as to question some of the hype that comes with the resurgence of popularity.

Prerequisites

The course is open to any students who have taken introductory probability/statistics, linear algebra, and calculus. Knowledge of basic ML concepts is preferred. Students must be competent in Python and NumPy, or be able to learn NumPy quickly.

Readings

The primary text will be Deep Learning by Ian Goodfellow, Yoshua Bengio, and Aaron Courville. The text is available online by chapter in html, but it should serve as a good reference and is worth purchasing. The following are also recommended as reference texts:

- Pattern Recognition and Machine Learning, Christopher Bishop
- Deng & Yu's monograph on Deep Learning: Methods and Applications

Geoffrey Hinton (Univ of Toronto and Google), one of the founders of modern DL, taught a Coursera class in 2012. It is a bit dated, but he gives such beautiful explanations and intuitions that his lectures are well worth viewing. Coursera has since deprecated the material, but it is available via YouTube. Many of his tutorials and invited talks are available on the web and are highly recommended.

CSCI 5922, Spring 2020

We will use Piazza for class discussion. Rather than directly emailing the instructors or TA, you are encouraged to post your questions to Piazza. You are also encouraged to respond to other students' questions on Piazza.

Grades

Semester grades will be based on the following:

- 40% Assignments
- 20% Midterm
- 5% Project proposal
- 30% Final project
- 5% Participation and attendance

Jan 13: Introduction

- Reading: Deep Learning, Chapter 1
- Slides
- Assignment #0 (Self Assessment)

Jan 15: History, perceptrons

- Reading: Deep Learning, Chapter 1
- Slides

Jan 22: Empirical Risk Minimization

- Reading: Deep Learning, Chapters 2-5 (review as needed)

Jan 27: Gradient descent and SGD

- Reading: Deep Learning, 6.0-6.2
- Combined instructor's notes for Jan 22 and Jan 27
- Assignment 1

- Reading: Deep learning, 6.3-6.6
- Additional reading
  - Automatic differentiation in machine learning: a survey
- Notebooks
- Slides

Feb 3: Backprop, CNNs

- Reading: Deep learning, 9.0-9.3
- Additional reading:
  - Bishop, Pattern Recognition and Machine Learning, Section 5.3
  - Karpathy's CS231n notes on CNN
- Slides

Feb 5: CNNs for image classification

- Reading: Deep learning, 9.4-9.11
- Additional reading:
  - AlexNet paper
  - Feature visualization
- MNIST notebook
- Assignment 2 - Backprop for MLPs
- Slides

Feb 10: Dual Encoder & Approximate Nearest Neighbor

- Additional Readings (all optional):
  - FaceNet paper
  - Book on Nearest Neighbor (no free pdf available)
  - Hierarchical Navigable Small World graph paper
  - ANN Benchmark
- Slides

CSCI 5922, Spring 2020

- Additional Readings (all optional):
  - Neural Language Model
  - SENNA
  - RNN Language Model/Linguistic Regularity
  - Word2vec (2013a, 2013b)
- Slides

Feb 17: Recurrent Neural Network

- Reading: Deep learning, 10.1-10.6
- Slides