CSE 253: Neural Networks for Pattern Recognition

Winter 2019 Syllabus

MWF: 2:00-2:50PM Center 109

OR

MWF: 3:00-3:50PM Center 109

Professor: Gary Cottrell

Office: CSE 4130; Phone: 858-534-6640

Gary's Office hours: Tuesdays noon-1, Wednesdays 4-4:50PM or by appointment

TAs:

Palash Agrawal	p6agrawa@ucsd.edu
Sai Aditya Chitturu	schittur@ucsd.edu
Tejash Desai	tvdesai@ucsd.edu
Paarth Neekhara	pneekhar@ucsd.edu
Anmol Popli	apopli@ucsd.edu
Yao Qin	yaq007@ucsd.edu
Jenny Hamer (tutor)	jhamer@ucsd.edu

Note: There will be no final; instead, there will be two in-class midterms. Instead of a final, we will have final project presentations on March 18th and 20th, (Monday and Wednesday of finals week) 3-6PM.

We will use Chris Bishop's 1995 book, *Neural networks for Pattern Recognition*, which can be found on the web as a pdf. This text contains a solid introduction to pattern recognition beyond just neural nets, especially the underlying statistical foundation. The text covers traditional pattern recognition, probability density estimation, single and multiple layer networks. Since the book is from 1995, well before deep networks became popular, I will add reading of primary sources. Many of the readings are already posted on Piazza.

Required coursework

The required work for this course will consist of two in-class midterms, and *five* programming assignments, at a brisk pace. Many of the programming assignments will include written homework. *Hence you will need to devote significant effort to this course*. Your programming assignments will count for 70% of your grade; the two midterms 15% each, and we will use clickers (5% extra credit, for participation only). Due to the inadequacies (i.e., bugs!) of the interface between the iClicker program and TritonEd, your clicker scores may not be available during the quarter (or, at least, may not make sense!). The programming assignments will be due as follows:

PA 1: (paired assignment, 10% of grade) logistic regression and softmax regression.

Handed out: January 7th; Due: January 17th.
PA 2: (paired assignment, 15% of grade) back propagation.
Handed out: January 16th; Due: January 30th.
PA 3: (groups of 3-4 people, 10% of grade): Convolutional networks.
Handed out: January 30th; Due: February 13th.
PA 4: (groups of 4-5 people, 15% of grade): LSTM (recurrent) networks.
Handed out: February 13th; Due: February 27th.
PA 5: (groups of 5+ people 20% of grade): *Final project of your choice*. See "Final Project Guidelines" on the resources page. Handed out: Monday, February 25th; Project proposal due Friday, March 1st; Progress report due March 8th. Draft final report due Friday, March 15th. Final report and presentation of project March 18th and 20th.

Within the 70% of your points based on the programming assignments, each assignment will be weighted based on the perceived difficulty of the assignment (see weightings above). For *all* group work, whether the programming assignment or the final project, *every member of the group must include a (short) paragraph on what their contribution to the project was in the report.* **NO EXCEPTIONS.**

Late Policy

Assignments can be turned in a day late with a 10% penalty. Two days late: 20%. Three days late: 50%. More than three days: 100%.

Course schedule with Readings (Chapter numbers refer to Bishop 1995, others as noted)

Week 1: Pattern Recognition Overview (chapter 1 of Bishop). Single-layer networks: Linear regression, logistic regression, and softmax regression (Parts of Chapters 3 and 6).

Week 2: Perceptrons (Chapter 3); Activation functions; Forward & Back propagation (Chapter 5); Representations (see Chapter8_with_notes_from_GWC, hinton86, Sejnowskispeech1987 under Readings).

Week 3: Tricks of the trade (see lecun_98_efficient_backprop.pdf in Readings, Batch_Normalization.pdf in Readings); objective functions (Chapter 6, Sections 6.1, 6.7, 6.9, chopra05.pdf & hadsell_chopra_lecun_2006.pdf in Readings); improving generalization.

Week 4: Convolutional Networks (see lecture notes, Chapter8_with_notes_from_GWC pp. 348-352, lecun89e in Readings).

Week 5: More Convnets; see Fergus slides online **in-class Midterm February 6th**

Week 6: Recurrent networks (Chapter8_with_notes_from_GWC pp. 354-361, see papers by Jordan, St. John, Elman under Readings); LSTM networks (will post Schmidhuber paper)

Week 7: Attention networks; more applications (see DRAM and neural turing machine paper in readings (remind me to post if not there)

Week 8: Hopfield Networks; Boltzmann machines; Restricted Boltzmann Machines (remind me to post papers if not there)

Week 9: Deep Reinforcement Learning; selected papers; **in-class Midterm March 6th,**

cumulative (emphasis on second half)

Week 10: Selected recent papers; Ethics of AI

First assignment

The first reading assignment is to read Chapters 1 & 2 for general background (we will not cover this in any detail, but it is good for you to know), and read Chapter 3 for background on the programming assignment. The first programming assignment will be posted on Piazza, and will include some written work.

Resources

We will endeavor to answer your questions on Piazza in a timely manner. We will hold office hours and/or sections almost every day of the week. There are many online deep learning courses that have good lectures available. If you are not getting answers to your questions on Piazza within 24 hours, please text the professor at 619-823-3033 and I will give the TAs holy hell.

- There are courses online where the giants in the field explain neural networks and deep learning. Geoff Hinton had a coursera neural network course, but it doesn't seem to be available to people who didn't sign up for it. He has a page of tutorials here. It looks like many of the lectures are available on Youtube if you search for "youtube hinton neural network lecture" – they are the ones that say "Lecture 1.1", etc.
- Andrew Ng has a coursera course on deep learning here. You can sign up for a 7-day free trial.
- Stanford's convolutional network course is here. The initial web page has a convolutional network running in your browser! The 2017 lectures are free and on youtube.
- Andrej Karpathy's lectures for that course from Winter 2016 are here.
- The neural network playground is a great tool to play around with a neural network and get insights into the kinds of features they learn. You can vary the number of hidden layers, the activation functions, the learning rates, etc.
- Our demo of face recognition is somewhat amusing, but requires matlab.
- The deep learning website has a LOT of resources…
- Andrej Karpathy's blog is just wonderful. There are lots of other blogs also.
- Two minute papers are an excellent way to keep up, without *really* keeping up…

Piazza

Hopefully, everything you need (aside from the above resources) will be posted on Piazza. You may pose questions there, and we encourage students to answer each other's questions. However, *don't post your code, either as part of a question or part of an answer.* Thanks!

Podcasts

The lectures are usually podcast unless there is a technical glitch. See podcast.ucsd.edu.

Missed exams: Exams may be made up within a reasonable time frame (e.g., with a day or so). Please advise the instructor if you know you will be out of town for an exam. We may give you an exam prior to your leaving town.

Grading policy

In graduate classes, I curve to a B+, unless doing so would lower your grade! My procedure is as follows: I take the median grade, add sufficient number of points to make it 88.5 (middle of a B+), add that number to all grades, and then round the result to the nearest whole number. I use, except at the top and the bottom, where I make the cutoff for an A+ 96 instead of 97, and at the bottom, I use Charles Elkan's dictum that "anyone who learns half of what I teach them deserves to pass."

Grading procedure

To "curve" to a B+, I use a linear technique: I take the median overall grade, and then add points to it to reach 88.5. This number of points is added to everyone's grade, and then round the result to the nearest whole number. Then I use standard cutoffs. After deciding the grades this way, I figure in extra credit. The extra credit is from using clickers. They are worth 5 final points. I take the maximum clicker grade minus 2 (i.e., drop 2 sessions), figure out your percentage of that, and apply it to the 5 points. This is then added to your score. 5 points is a significant amount, and can change your grade from, say, a B+ to an A. So get and register a clicker!

Academic Integrity

Working in small groups in programming assignments is required. However, on written homeworks (derivations and such), please follow the **Gilligan's Island rule**:

> (Dymond, 1986): No notes can be made during a discussion, and you must watch one hour of Gilligan's Island or some equally insipid TV show before writing anything down. Suspected cheating has been reported to the Dean in the past, and will be again.

On midterms, of course, please do not bring notes to yourself up your sleeve, or read off of other students' tests, etc. We have had complaints of cheating in the past during tests with graduate students. It seems unbelievable, but it happens. We will be watching closely. Anyone who is caught cheating on a test will receive a 0 for that test. We will give out two versions of each test.

To detect instances of academic integrity violations in programming assignments we will use 3rd party software. We recommend you do not put your name in your code, unless you don't mind this potential exposure. Including your name and/or PID will disclose that information to the 3rd party.

Please sign the integrity pledge below and turn it in.

Prerequisites

This course has no formal prerequisites, but some mathematical sophistication is required. You should know some probability, and know how to sample from a distribution. You should know linear algebra, and vector calculus (at least, partial derivatives). This course is intended for CSE, Cognitive Science and ECE graduate students. The use of numPy is strongly encouraged for the first two assignments. After those two, we will switch to using a deep learning platform – we plan to use PyTorch because it has good facilities for inspecting internal variables, which can aid in debugging.

Texts

Neural Networks for Pattern Recognition, by Chris Bishop, Oxford University Press. Again, you can find this in pdf online

The Deep Learning Book. This is a new book that came out last year, and is again, free online. It's a little rough around the edges.

Supplementary reading:

Duda, Hart and Stork *Pattern Classification* (2nd Ed). Wiley

For the Physics inclined:

Introduction to the Theory of Neural Computation. Hertz, Krogh & Palmer, Addison Wesley.

For the Cognitively inclined:

Parallel Distributed Processing, Vols 1 & 2, edited by Rumelhart and McClelland (MIT Press, 1986) and *Explorations in PDP*, by McClelland & Rumelhart (MIT Press, 1987) (This is a good way to get a working introduction to neural nets).

Diversity and Inclusion

We are committed to fostering a learning environment for this course that supports a diversity of thoughts, perspectives and experiences, and respects your identities (including race, ethnicity, heritage, gender, sex, class, sexuality, religion, ability, age, educational background, etc.). Our goal is to create a diverse and inclusive learning environment where all students feel comfortable and can thrive.

Our instructional staff will make a concerted effort to be welcoming and inclusive to the wide diversity of students in this course. If you have a particular gender pronoun you prefer, or some other way we can make you feel more included please let one of the course staff know, either in person, via email/discussion board, or even in a note under the door. Our learning about diverse perspectives and identities is an ongoing process, and we welcome your perspectives and input.

We also expect that you, as a student in this course, will honor and respect your classmates, abiding by the UCSD Principles of Community (https://ucsd.edu/about/principles.html). Please understand that others' backgrounds, perspectives and experiences may be different than your own, and help us to build an environment where everyone is respected and feels comfortable. In particular, be respectful in responding to other students on the Piazza page.

If you experience any sort of harassment or discrimination, please contact the instructor as soon as possible. If you prefer to speak with someone outside of the course, please contact the Office of Prevention of Harassment and Discrimination: https://ophd.ucsd.edu/.

Basic Needs Resources

Are you eating properly? Do you have adequate access to nutritious food? Do you have stable housing? Are you homeless or couch surfing? If you or someone you know has food and/or housing insecurity, please note: http://basicneeds.ucsd.edu

The Triton Food Pantry (in the old Student Center), https://www.facebook.com/tritonfoodpantry/ is free and anonymous, and includes produce.

Financial aid resources, the possibility of emergency grant funding, and off-campus housing referral resources are available.

CAPS and college deans can connect students to the above resources, as well as other community resources and support.

Excel with Integrity Pledge

The *Excel with Integrity pledge* affirms the UC San Diego commitment to excel with integrity, both on and off campus, in academic, professional, and research endeavors.

According to the International Center for Academic Integrity, academic integrity means having the courage to uphold honesty, fairness, responsibility, respect & trust even when difficult.
Creating work with integrity is important because otherwise we are misrepresenting our knowledge and abilities and the University is falsely certifying our accomplishments.
And when this happens, the UCSD degree loses its value and we've all wasted our time and talents!

Student Name: ____

The Student's Excel with Integrity Pledge:

I am **fair** to my classmates and instructors by not using any unauthorized aids.
I **respect** myself and my university by upholding educational and evaluative goals.
I am **honest** in my representations of myself and of my work.
I accept **responsibility** for ensuring my actions are in accord with academic integrity.
I show that I am **trustworthy** even when no one is watching.

Please affirm your adherence to this pledge by *writing below* the following statement:
***I Excel with Integrity*.**

Signature: ________________________________ **Date:** ___________