Department of Computer Science, University of Illinois at Chicago Spring 2020

CS 512 — Advanced Machine Learning (CRN: 42204)

Course Syllabus

This course will contribute to your course requirement as a standard 500-level course.

For PhD students, it can be used to fulfill your WCP requirement.

Time: TR 12:30 – 1:45 PM

Classroom: Stevenson Hall 319

URL: via Blackboard / Piazza (https://piazza.com/class/k4i1mw5d3rwrl)

Instructor: Xinhua Zhang

Office: North End, Level 3, Richard Daley Library

Phone: 312.413.2416

E-mail: zhangx@uic.edu (preferred)

Office hour: 2-3 PM **Tuesday**. Contact me by email if you would like to meet at other time.

TA: Yingyi Ma (yma36@uic.edu)

Office hour: 11am - 12pm **Thursday**. Office: **TBD**

Prerequisites

MATH 310/320, STAT 401, CS 251, CS 412; or consent of the instructor.

For graduate students, these prerequisites are only advisory.

A self-evaluation quiz with solution is available on Piazza and Blackboard, along with some math background readings. Do it yourself and you do not need to submit it.

Course Goals:

- Students will be able to have an in-depth understanding of the principle and characteristics of advanced machine learning task settings (e.g., structured prediction, convex optimization, deep learning for complex data).
- Students will be able to scale machine learning techniques to big datasets, by leveraging new structures in the data and new computational tools that emerge even after the completion of the course.

- Students will be able develop and analyze novel problem formulations and machine learning techniques that adapt to data analysis problems emerging in new applications.

Restrictions

Restricted to students in the following colleges/schools: Engineering or Graduate College. Students who have taken CS 594 (Advanced Machine Learning) in **Fall 2016 or Spring 2018** are NOT eligible for registration.

Credit Hours

4 graduate hours

Textbooks (Required, all free)

[**BV**] Stephen Boyd and Lieven Vandenberghe. Convex Optimization. Cambridge University Press, 2004. PDF available at https://web.stanford.edu/~boyd/cvxbook

[**Jordan**] Michael I. Jordan. Graphical Models. Unpublished lecture notes (Ch 2, 3, 4, 9, 10, 11). PDF available on Blackboard

[**Murphy**] Kevin P. Murphy. Machine Learning: A Probabilistic Perspective. MIT Press, 2012. Free book available electronically via UIC library

[**GBC**] Ian Goodfellow, Yoshua Bengio, and Aaron Courville, Deep Learning. MIT Press, 2016.

Free book at: http://www.deeplearningbook.org Single PDF (slightly different from the online version) for download (65 MB): https://uofi.box.com/s/nv6idmcb5w9pyqay4966cxwdaaz547jy

We will also use recent research papers and excerpts of relevant background material from available textbooks, and supplemental notes for specific topics.

Matlab is free on Webstore: https://webstore.illinois.edu/home/NewsItem.aspx?PostID=329.

Electronic Communication

1. Blackboard will contain all materials relevant to the class, syllabus, lecture notes, tutorials, assignments and solutions, grades, etc. You can also see you own grades. For the first two weeks when registration is not finalized, course materials will also be made available on Piazza.

2. For general announcements and notifications, I will send email to the whole class via Blackboard. Please check your email frequently, especially around deadlines (homework and exams).

3. Piazza will be used for Question and Answer ONLY. The system is highly catered to

getting you help fast and efficiently from classmates and myself. Rather than emailing questions to the instructor, you are encouraged to post on Piazza your *technical* questions, general questions about the course content, assignments, grading rubric, etc. If you have any problems or feedback for the Piazza developers, email team@piazza.com.

Find our Piazza class page at: https://piazza.com/class/k4i1mw5d3rwrl

4. If you have any personal or non-technical questions such as medical/grievance considerations, please send an email directly to the instructor (zhangx@uic.edu) or the TA (if instructed to do so, e.g. dispute of grading).

Evaluation Criteria

	%	**tentative** deadlines	Note
Assignment 1	13 %	Week 5	Some written questions for individual work, and a project on conditional random fields with inference and learning (free language, utility code will be provided in Matlab & Python)
Assignment 2	13 %	Week 8	A project on deep learning using PyTorch, focusing on convolution neural network and image processing
Mid-term (75 min in class)	16 %	Week 9	Basic deep learning and graphical model. You can bring one cheat sheet.
Assignment 3	13 %	Week 13	A project on adversarial learning with recurrent neural networks
Convex optimization			Work on sample questions in preparation of the exams
Course Project	20 %	Week 16	Deep learning project (open, free language) **Oral presentation in week 15** (?? minutes per group)
Final exam	25 %		**Cumulative and comprehensive** (one cheat sheet)

Assignments and course project are in groups of 3-4 students. We will make 10 groups. Assignments 1-3 aim to develop students' ability to apply machine learning tools to different data analysis problems and perform thorough experimentation, benchmarking, and empirical analysis. Tasks will be specified in Assignments 1-3. Assignment 1 contains some written questions for individual work.

Midterm and final exams are individual. They will assess students' understanding of different machine learning tasks and techniques, especially their theoretical properties and underlying principles.

Course projects, including reports and presentations, aim to develop students' ability to create novel machine learning techniques for different real-data analysis problems by leveraging their specific structures and computational resources at hand.

Grading Policy. All the evaluations will be graded out of 100, and their weighted average will be used to determine the final letter grade (A/B/…) **based on threshold**:

A: 80 - 100
B: 60 ~ 79
C: 40 ~ 59
D: 0 ~ 39

I reserve the right to **raise** your letter grade.

Detailed notes:

1. The mid-term will be **75 minutes** with **one letter-sized cheat sheet (both sides)**. It will be given in the classroom during class time. Therefore, **no make-ups** will be given. Partial grading will be used.

   The mid-term will cover graphical models and basic deep learning (tentatively). Sample questions will be given.

2. In Assignment 1 you will build a classifier which recognizes “words" from images. This is a great opportunity to pick up practical experiences that are crucial for successfully applying machine learning to real world problems, and evaluating their performance with comparison to other methods. In particular, you will implement probabilistic inference algorithms such as message passing, train a CRF model using off-the-shelf solvers, compare with max-margin methods, and test the performance under transformations. Although you may use any language of your choice, we will only provide utility code in Matlab and Python. Grading will be based on your results. Some written questions are given for individual work.

3. In Assignment 2, your work in Assignment 1 will be extended to deep learning. You will reimplement CRF using the modules in PyTorch, add a convolution layer implemented by yourself, and test stochastic optimization with approximate inference based on sampling. You can use GPU on AWS or Google colab.

   PyTorch must be used for this project. Grading will be based on your results and your code.

4. In Assignment 3, you will work on an advanced topic of deep learning: virtual adversarial training on recurrent neural networks, applied to semi-supervised text classification. You will implement the algorithm in PyTorch, and benchmark several competing training algorithms and deep architectures.

5. The course project will be on deep learning, using the AWS platform. You will have the opportunity to design your own project. We plan to form 10 groups, and each group will have 15 minutes to present their projects orally in the two sessions on week 15 (12 min talk and 3 min Q&A). Only *initial* results will be needed for oral presentation, and the detailed results can be submitted in the final report on Monday of week 17.

6. The final exam will be **120 minutes** with **one letter-sized cheat sheet (both sides)**. It will be comprehensive.

Course Policies

Academic dishonesty will not be tolerated. Please see the CS department policy below on the topic; this policy specifies penalties for violations. Academic misconduct will be handled according to UIC's Student Disciplinary Policy: http://dos.uic.edu/conductforstudents.shtml

What is academic dishonesty? To hand in any work which is not 100% the student's creation, unless you are explicitly allowed to do so.

Collaboration Policy

1. **Homework and course project**: all members of each group should make nearly the same level of contribution to each project. So in a group of four, each member contributes 1/4 to Assignment 1, 1/4 to Assignment 2, 1/4 to Assignment 3, and 1/4 to the course project. It is not allowed that one works on Assignment 1 alone, one on Assignment 2 alone, etc. All members should be responsible for the whole submission of the team, not only his/her own contributed part. You are supposed to understand the work of your teammates inside out, and be able to answer questions when asked. If one member plagiarized, then all members of the team will receive the same penalty.

   No team may give other team any portion of their solutions or code, through any means. Students are not allowed to show each other any portions of code or homework, unless they are on the same team.

   Discussion of homework assignments and solutions with other students (including outside the group) is permitted, including all assignments and the course project. However, for the written questions in Assignment 1, each student must

   1) submit his or her own write-up for the homework assignment and fully understand what he or she submits;

   and

   2) be prepared to explain his or her homework assignment submissions to the instructor and teaching assistant if his or her "full understanding" is in doubt.

2. **Exams (midterm and final): All work on both exams must be individually performed.**

Policy for Missed or Late Evaluation

1. **Late submission**s: Late submissions will not be accepted in any case, unless there is a **documented** personal emergency. Arrangements must be made with the instructor as soon as possible after the emergency arises, preferably well before the homework due date.

   Advice: If for whatever reason you don't manage to finish an assignment, hand in what you have. Partial credit will be given.

2. **Statute of limitations**: **Three weeks!** No grading questions or complaints — no matter how justified — will be listened to **three** weeks after the item in question has been returned.

3. **Missed exams**: Missed exams cannot be made up unless there are extenuating circumstances

(death, severe illness, etc.) **and** the student has e-mailed instructor **in writing** **before** the exam. A mark of 0 is earned for the exam if the above policy is not adhered to.

Religious Holidays

Students who wish to observe their religious holidays shall notify the faculty member by the tenth day of the semester of the date when they will be absent. The faculty member shall make every reasonable effort to honor the request, not penalize the student for missing the class, and if an examination or project is due during the absence, give the student an exam or assignment equivalent to the one completed by those students in attendance. If the student feels aggrieved, he/she may request remedy through the campus grievance procedure. http://oae.uic.edu/docs/ReligiousHolidaysFY20152017.pdf

CS department policy on academic dishonesty

The CS Department will not tolerate cheating by its students. The MINIMUM penalty for any student found cheating will be to receive an F for the course and to have the event recorded in a department and/or College record. The maximum penalty will be expulsion from the University.
Cheating includes all the following, though this is not a complete list:

- Copying or any other form of getting or giving assistance from another student during any test, quiz, exam, midterm, etc.
- Plagiarism–turning in writing that is copied from some other source.
- Obtaining solutions to homework by posting to the Internet for assistance, purchasing assistance, obtaining copies of solutions manuals for instructors, and obtaining copies of previous year's homework solutions.
- Computer programs: Any time you look at another student's code, it is cheating. (Exception: If you are EXPLICITLY told that you may do so by the instructor.)

For computer programs, if for some reason we cannot determine who copied from whom, we may, at our discretion, give failing grades to both students.

It is the responsibility of all engineering and computer science professionals to safeguard their company's "trade secrets." An employee who allows trade secrets to be obtained by competitors will almost certainly be fired. So, YOU are responsible for making sure that your directories have permissions set so that only you can read your files, for being sure to log out at the end of working in the computer lab, etc.

A (Very) Tentative weekly schedule

Week	Topic	Readings	Assessment
1	Course introduction and feedforward networks	Chapter 6 of GBC	
1	Feedforward networks and back-propagation	Chapter 6 of GBC	Tutorial on deep learning out
2	Convolutional network	Chapter 9 of GBC	

2	Recurrent network	Chapter 10 of GBC	
3	Probability and graph basics AWS introduction	Chapter 2 and 10.1.1-10.1.4 of Murphy, Tutorial on AWS	Assignment 1 out
3	Directed graphical models (Bayes nets) and Undirected graphical models	Chapter 2 of Jordan	
4	Max-likelihood learning on GM and Conditional Random Fields (CRFs)	Chapter 8 of Jordan (excluding section 8.3.4) Chapter 19.6 of Murphy (optional 9.3-9.4 of Murphy)	Tutorial on graphical models out
4	Variable elimination	Chapter 3 of Jordan Chapter 19.1-19.4 of Murphy	
5	Sum-Product algorithm (belief propagation)	Chapter 4 of Jordan	
5	Sum-Product algorithm (belief propagation)	Chapter 4 of Jordan	Assignment 1 due
6	NN regularization	Chapter 7 of GBC	Assignment 2 out
6	Autoencoder and restricted Boltzmann machines	Chapter 14 of GBC	
7	Adversarial learning	Papers and tutorials	
7	Convex sets	Chapter 2 of BV	
8	Convex sets	Chapter 2 of BV	Tutorial on convex optimization out
8	Convex function	Chapter 3 of BV	Assignment 2 due
9	Convex function	Chapter 3 of BV	
9	**Midterm (covering everything up to adversarial learning)**		
10	Convex problems	Chapter 4 of BV	Assignment 3 out
10	Convex problems	Chapter 4 of BV	Project proposal due (Sunday)
	Spring break		
11	Convex and nonconvex optimization (for neural networks)	Chapter 9.1-9.5 of BV Chapter 8 of GBC	
11	Convex and nonconvex optimization (for neural networks)	Chapter 9.1-9.5 of BV Chapter 8 of GBC	

12	Duality	Chapter 5 of BV	
12	Duality	Chapter 5 of BV	Assignment 3 due
13	Mixture model and EM algorithm	Chapter 9 and 10 of Jordan	
13	Variational autoencoder + GAN	Papers and tutorials	
14	Deep reinforcement learning	Papers and tutorials	
15	Course project presentation		