COS 324 – Introduction to Machine Learning

Ryan P. Adams
Princeton University

Instructor:	Prof. Ryan P. Adams (OH Mon and Wed 3-4pm in CS 411)
Teaching Assistants:	Jad Rahme (OH Tue 6-8pm in Fine Hall 216)
	Farhan Damani (OH Mon 7-9pm outside CS 242)
	Fanghong Dong (OH Wed 4-6pm in CS 2nd floor tea room)
Lectures:	Monday and Wednesday, 1:30-2:50pm
Location:	CS 104
Precepts:	Wednesday 7:30-8:20pm in Friend Center 009
	Friday 1:30-2:20pm in Andlinger Center 017
	Friday 2:30-3:20pm in Andlinger Center 017
URL:	`https://www.cs.princeton.edu/courses/archive/spring19/cos324/`
Contact:	`cos324-s19@lists.cs.princeton.edu`

Course Description

This course provides a broad introduction to machine learning, probabilistic reasoning and decision making in uncertain environments. The course should be of interest to undergraduate students in computer science, applied mathematics, sciences and engineering, and lower-level graduate students looking to gain an introduction to the tools of machine learning and probabilistic reasoning with applications to data-intensive problems in the applied sciences, natural sciences and social sciences.

For students with interests in the fundamentals of machine learning and probabilistic artificial intelligence, this course will address three central, related questions in the design and engineering of intelligent systems. How can a system process its perceptual inputs in order to obtain a reasonable picture of the world? How can we build programs that learn from experience? How can we design systems to deal with the inherent uncertainty in the real world?

Our approach to these questions will be both theoretical and practical. We will develop a mathematical underpinning for the methods of machine learning and probabilistic reasoning. We will look at a variety of successful algorithms and applications. We will also discuss the motivations behind the algorithms, and the properties that determine whether or not they will work well for a particular task.

Course Website

The course URL is `https://www.cs.princeton.edu/courses/archive/spring19/cos324/`. We'll also use Piazza at `http://piazza.com/princeton/spring2019/cos324`. Most questions about the course, lecture/precept material, or the assignments should be addressed via Piazza. The course instructors will regularly check this discussion board with the goal of posting responses within 24 hours. Students taking the class are also encouraged to post responses. Code examples can be posted, but don't post anything

you wouldn't be expected to share with other students in the class as per the collaboration policy. Long, detailed questions are probably best answered during office hours. Questions that are not appropriate for the discussion board may be sent to the staff via email. Use your judgement.

Precepts

There will be a recommended hour-long weekly section led by the teaching assistants. You are welcome to attend whichever precept fits your schedule, assuming there is space. Precepts will be interactive and contain discussion of questions and examples related to the course. They will include discussion of homework assignments, and worked out examples of questions that might be good preparation for the exams.

Textbook and Course Materials

There is no required textbook for the course. This course has its own notes that are considered the required reading. Nevertheless, people learn in different ways and seeing the material presented in different formats can be valuable. To that end, additional optional material is linked on the course website and several books provide useful additional reading:

- Kevin Murphy. *Machine Learning: A Probabilistic Perspective*. MIT Press. 2012.
- Christopher M. Bishop. *Pattern Recognition and Machine Learning*. Springer. 2011.
- David J.C. MacKay. *Information Theory, Inference, and Learning Algorithms*. Cambridge University Press. 2003. Freely available online at `http://www.inference.org.uk/itila/book.html`.
- Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning*. Springer. 2001. Freely available online at `http://www-stat.stanford.edu/~tibs/ElemStatLearn/`
- Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. *An Introduction to Statistical Learning*. Springer. 2013. Freely available online at `http://www-bcf.usc.edu/~gareth/ISL/`
- Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning: An Introduction*. MIT Press. 1998. Freely available online at `http://incompleteideas.net/book/the-book-2nd.html`

Prerequisites

Students should be comfortable with writing non-trivial programs in Python. Students should have a background in basic probability theory, and some level of mathematical sophistication, including calculus and linear algebra.

Requirements and Grading

There are six homework assignments, each representing 10% of the final grade. There is a midterm and a final, each representing a further 20% of the final grade.

Assignments

There are six homework assignments, each representing 10% of your total grade. These homework assignments will involve components that are theoretical or conceptual, as well as some practical implementation. Each assignment will have an associated URL to the CS dropbox for uploading. It is recommended that assignments be completed in LaTeXand compiled to PDF. Homework assignments should be done individually.

Collaboration Policy

We want you to be able to discuss the class material with each other, but we want the homework you submit to be your own work. More specifically:

- You may never:
  - Share code.
  - Share writeups.
- You may always:
  - Discuss the related concepts and the high-level approach.
  - Discuss the results of your experiments at a high level, e.g., "I got 90% test accuracy."
- You should be wary of discussing details of proofs, your code, or results at an implementation level, rather than at the "big idea" level.
- In your assignment writeup, state who you discussed the problems with.
- It is prohibited to search the internet for assignment solutions.

Late Policy

Homework assignments may be turned in up to a week late for a 50% penalty. There will be no exceptions and no further extensions. Plan ahead.

Changelog

- 2 February 2019 – Updated to spring 2019.
- 13 September 2018 – Initial version.