
ISyE 6740 - Computational Data Analysis / Machine Learning I Spring 2018 Tentative Syllabus

Machine learning is a field of computer science that gives computers the ability to learn without being explicitly programmed. The course is designed to answer the most fundamental questions about machine learning: What are the most important methods to know about, and why? How can we answer the question 'is this method better than that one' using asymptotic theory? How can we answer the question 'is this method better than that one' for a specific dataset of interest? What can we say about the errors our method will make on future data? What's the 'right' objective function? What does it mean to be statistically rigorous?

This course is designed to give graduate students a thorough grounding in the methods, theory, mathematics and algorithms needed to do research and applications in machine learning. The course covers topics from machine learning, classical statistics, and data mining. Students entering the class with a pre-existing working knowledge of probability, statistics and algorithms will be at an advantage, but the class has been designed so that anyone with a strong numerate background can catch up and fully participate. Some experience with coding are expected (at a language of your choice, e.g., MATLAB or Python.)

For detailed course topics, please see the tentative course schedule.

Class Time and Location: MWF, 10:10-11:00am, 2018, Monday Jan. 9 - Monday April 23 (Final Instructional Class). Manufacturing Related Disciplines Complex (MRDC), Room 2404.

Instructor: Prof. Yao Xie, Groseclose #339, email: yao.xie@isye.gatech.edu

Instructor Office Hour: Wednesday 11am-noon, Groseclose 339

Class TA:

Lead TA: Haoming Jiang, jianghm@gatech.edu

TA: Shuang Li, sli370@gatech.edu

TA Office Hour: Tuesday, 1:30-2:30pm, Main Building 224-GTA Office Hour Room.

Class Website: T-square

All class material available on our website includes

- Announcements
- Course syllabus
- Homework assignments and solutions
- Slides and other lecture material
- Practical exams
- Your course grades on exams and homework
- Any important announcements

Make sure the scores in T-square are consistent with what you got. We will not make any change in grading for works older than 2 weeks.

Class Mailing List: Registered students are automatically subscribed to the class mailing list.

Textbook: the course material will be based on lectures and slides posted on T-square.

References:

(PRML) *Pattern recognition and Machine Learning*, Christopher M. Bishop.

(ESL) *The elements of Statistical Learning: Data Mining, Inference, and Predictions*, 2nd edition, Trevor Hastie, Robert Tibshirani, and Jerome Friedman.

Prerequisites: undergraduate level basic probability, linear algebra, and statistics.

Honor Code: For any question involving Academic Honor Code issues, please consult www.honor.gatech.edu

Programming language: Any. MATLAB, R, Python... at your choice.

Grading Policy: Homework - 25%, Midterm 1 - 20%, Midterm 2 - 20%, Final - 35%.

Homework: Homework should be submitted before the deadline set in T-Square. No late submission will be accepted through email. We strongly encourage to use LaTeX for your submission. Assignments will include both exercises and computer problems; the computer problems will ask you to carry out statistical analysis using computer statistical software. Keep in mind that you should not hand in raw computer output. Conclusions and interpretation of results are more important than good printouts. You can work together with other students on homework, as long as you write up and turn in your own solutions. You are also allowed (and encouraged) to ask me questions, although you should try to think about the problems before asking. Request for re-grading the Homework/Exams/Quizzes should be made within a week of returning Homework/Exams/Quizzes. Any kind of academic misconduct is subject to F grade as well as reporting to the Dean of students.

Exams: There will be two in-class midterm exams and one final exam. The exams are open-book. You are not allowed to use your cell phone or computer. If there is any conflict of time, please let us know beforehand. There are **no make-ups**.

Midterm 1: Monday February 12, 2018, in class.

Midterm 2: Friday March 16, 2018, in class.

Final Exam Date: Wednesday May 2, 2018, 8:00-10:50am. Final exam is comprehensive and an emphasis on material after Midterms 1 and 2.