Syllabus, CMSE 820

Mathematical Foundations of Data Science

Spring 2017

Course Description: The ability to process, extract, and utilize insightful information from large amounts of data has become a desired, if not necessary, skill in almost every field of industry and science. Among other benefits, such information can provide useful knowledge, support decision-making, uncover hidden trends, and enable deeper understanding of observed phenomena. This course will cover some of the main problems and challenges encountered in data science. It will introduce you to the fundamental mathematical principles of data science that underlie the algorithms, processes, methods, and data-centric thinking.

Instructor Information

- **instructor:** Matthew Hirn
- **office:** 2507F, Engineering Building
- **email:** mhirn@msu.edu
- **phone:** (517) 432-0611
- **course webpage:** MSU Desire2Learn (D2L) course page

Meeting Time and Classroom:

- Tuesday, Thursday 10:20 AM – 11:40 AM
- 202 Urban Plan & Land Arch Bldg

Office Hours:

- Tuesday, 3:00 - 4:00 PM
- Friday, 3:00 - 4:00 PM

Midterm Exam:

- Thursday, March 2, 10:20 AM – 11:40 AM
- 202 Urban Plan & Land Arch Bldg

Final Exam:

- Thursday, May 4, 7:45 AM – 9:45 AM
- 202 Urban Plan & Land Arch Bldg
- This is a qualifying exam course for the CMSE PhD. Therefore the final exam is the qualifying exam for those students seeking a CMSE PhD!

Prerequisites:

- CMSE 802 or equivalent experience
- Differential equations at the level of MTH 235/255H/340+442/347H+442
- Linear algebra at the level of MTH 390/317H
- Probability and statistics at the level of STT 231

Resources (none required):

- *Learning with Kernels*, by Schlkopf and Smola.
  Available online through the MSU library.
- *Foundations of Machine Learning*, by Mohri, Rostamizadeh and Talwalkar.
  Available online through the MSU library.
- *Topics in Mathematics of Data Science*, MIT open course by Afonso Bandeira.
  Available online.
- More may be added later.

Grading:

- Homework Exercises: 35%
- Midterm: 15%
- Project: 15%
- Final Exam: 35%

How the course will work:

There is **no textbook** for the course. Lectures will be taken from the above resources, as well as other sources to be named later. After each lecture, I will post the lecture notes from class on the D2L website.

Exercises will be posted on D2L on a rolling basis, and each exercise will be scored out of 10 points. After each class there could be anywhere from zero to several new exercises; generally they will be due one week after they are posted. All exercises must be turned in through the D2L dropbox system. Exercises will be a mix of programming (in MATLAB) and "pen and paper" mathematical proofs. All solutions must be typed up and submitted as a pdf (ideally in LaTeX and compiled to generate a pdf). Each person is required to type and submit their own exercises, although you may discuss the exercises with your classmates.

Everyone will be required to develop and submit a **course project**. Even though this course is on the mathematical foundations of data science, we should not forget that this work is in service of solving applied problems. The project is an opportunity for you to explore an application domain of the theory we will develop.

The **midterm exam** will cover everything we have covered up to that point. The **final exam** will cover the entire course. These are "closed book" exams; no notes, computers, phones, or other aides are permitted during the exam.

Academic Honesty: Cheating in any form will not be tolerated and will be reported. You will receive a zero on any assignment in which there is a case of cheating. This includes, but is not limited to, plagiarism, failure to give proper citations, and copying another's work.

If you are preparing an assignment and have a question about whether you are adhering to this policy, please ask me. If you work on an assignment with other students, you must give credit to your collaborators.

MSU's policy on academic integrity can be found at the following URL:
`https://msu.edu/unit/ombud/academic-integrity/`

Disability Services: Accommodations for persons with disabilities can and will be made in this course. All arrangements will be organized through the RCPD office as MSU. Persons with disabilities who are interested in the available services should contact the MSU Resource Center for Persons with Disabilities (RCPD) at (517) 884-7273 or online at `http://www.rcpd.msu.edu`.