IEMS 351: Optimization Methods in Data Science

Instructor: Hao-Jun Michael Shi
TA: Shenyinying (Ruby) Tu
Lectures: MWF 12:00 – 12:50 CDT
Lab: W 11:00 – 11:50 CDT

Course Description: This course offers an introduction to nonlinear mathematical optimization with applications in data science. The theoretical foundation and the fundamental algorithms for nonlinear optimization are studied and applied to supervised learning models, including nonlinear regression, logistic regression, support vector machines, and deep neural networks. Students write their own implementation of the algorithms in the Python programming language and explore their performance on realistic data sets.

Prerequisites: IEMS 303 and 313; some programming course

Textbook: No textbook, but will be drawing significantly from the following references:

- *Ronald L. Rardin. "Optimization in Operations Research", first or second edition.*
- *Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. "An Introduction to Statistical Learning with Applications in R".*
- *Michael T. Heat. "Scientific Computing"*, *second edition*.
- *Stephen Boyd and Lieven Vandenberghe. "Introduction to Applied Linear Algebra – Vectors, Matrices, and Least Squares".*
- *Stephen Wright and Jorge Nocedal. "Numerical Optimization".*
- *Ian Goodfellow, Yoshua Bengio, and Aaron Courville. "Introduction to Deep Learning".*

Course Objectives:

- Students will be able to mathematically characterize optimal solutions for nonlinear optimization models
- Students will know the fundamental algorithms for unconstrained optimization
- Students will know how to specialize these algorithms for data science problems
- Students will be able to implement these algorithms in Python
- Students will be familiar with nonlinear regression, logistic regression, deep neural network, and support vector machine models
- Students will have practical experience in training these models using their own Python implementations
- Students will be able to write well documented and structured Python programs

Topics:

1. Unconstrained nonlinear optimization theory
2. Linear and nonlinear regression, logistic regression
3. Support vector machines
4. Numerical solution of linear systems

5. One-dimensional search algorithms (golden section search, bisection method)
6. Gradient descent
7. Newton's method
8. Gauss-Newton and Levenberg-Marquardt methods
9. Coordinate descent methods
10. Stochastic gradient method
11. Deep neural networks
12. Python programming (throughout the course)

Grading:

- Homework (30%)
- Programming Assignments (30%)
- Final Exam (40%)

Because undergraduate courses this quarter have been instituted to give Pass/No Pass grading, a grade of D or higher is required in order to pass the class. For the purposes of this course, this is obtaining 65% or higher in the class.

Remote Resources: As we all know, this quarter will be challenging due to the pandemic we are in. In order to bridge the gap between the remote and offline classroom setting, I want to draw attention to some resources that will be provided:

- Students are strongly encouraged to attend the live lecture held during class time, which will be held using Zoom. This allows for direct interaction between the students and instructor as in a normal classroom setting. That being said, I recognize that some students may have difficulty attending all of the lectures due to conflicts/time differences. For this reason, *lectures will be recorded and posted after class for later reference*.
- We will use a combination of slides, demos, and whiteboard notes during class.
- A PDF version of the slides will be provided before lectures.
- Please use online discussion boards on Piazza for questions regarding the homework or class material. Questions will be answered between 9 AM to 7 PM from Monday through Friday. Please allow one "business day" for a response.
- We will schedule multiple online study hours to encourage collaboration between students. This will be set up as online Zoom meetings that are open to anyone during that time period.
- Google/Stack Overflow is your friend when it comes to programming questions.

Office Hours: Online office hours will be held weekly by both the instructor and TA. Scheduling will be determined in-class during the first class based on student availabilities. Both the instructor and TA will also be available by appointment.

Homework Policy: We will be having weekly written assignments due each Friday (at 11:59 PM) starting the 2nd week of the quarter. The written assignment with the lowest score will be dropped. We will also have a set of programming assignments that will be provided two weeks prior to the due date. (Advice: Start early!) No late submissions (both written and programming) will be accepted except under extraneous circumstances.

You are strongly encouraged to work together and discuss homework problems with your peers, but you must write down your own solution/code. We will provide optional online study hours over Zoom to encourage collaboration.

For regrades, please submit a written request over email. Regrades will only be accepted at the latest a week after homework is returned.

Labs: Lab sessions will discuss useful topics around numerical computation and Python programming. The first lab will be for help installing Anaconda 3 and going through the Python tutorial. The final lab will consist of a PyTorch tutorial.

Academic Integrity: See https://www.northwestern.edu/provost/policies/academic-integrity/academic-integrity-guide-september-2019.pdf. Do not copy from others (including programming), use material from previous classes, or share class materials.

Using Zoom: You are strongly encouraged to log in to your Northwestern Zoom account using your NetID and password. You can then create your Zoom profile, including a picture that will display when your Zoom video is off.

Q&A in Piazza: IEMS courses, and the IEMS Undergraduate Program, will be using Piazza to facilitate class Q&A and discussion. You can access Piazza from the course navigation at left. Further guidance about effective and appropriate use of Piazza will be posted here soon.

Class Recordings: This class or portions of this class will be recorded by the instructor for educational purposes. These recordings will be shared only with students enrolled in the course and will be deleted at the end of the end of the Spring Quarter, 2020 course. Your instructor will communicate how you can access the recordings.

Unauthorized student recording of classroom or other academic activities (including advising sessions or office hours) is prohibited. Unauthorized recording is unethical and may also be a violation of University policy and state law. Students requesting the use of assistive technology as an accommodation should contact AccessibleNU. Unauthorized use of classroom recordings – including distributing or posting them – is also prohibited. Under the University's Copyright Policy, faculty own the copyright to instructional materials – including those resources created specifically for the purposes of instruction, such as syllabi, lectures and lecture notes, and presentations. Students cannot copy, reproduce, display, or distribute these materials. Students who engage in unauthorized recording, unauthorized use of a recording, or

unauthorized distribution of instructional materials will be referred to the appropriate University office for follow-up.

Student Resources:

Academic Support and Learning Advancement, Learning During Covid-19:
https://www.northwestern.edu/academic-support-learning/academic-strategies/learning-during-covid-19.html

NU Student Resources for Remote Learning:
https://digitallearning.northwestern.edu/remote/students

Using Zoom, for Students:
https://www.youtube.com/watch?v=wbnyQwsVbiY

Northwestern Financial Aid Emergency Funding:
https://undergradaid.northwestern.edu/types-of-aid/covid-19-requests.html

CAPS Covid-19 Resources:
https://www.northwestern.edu/counseling/outreach-education/covid-19-resources/index.html
Note that CAPS will continue to provide services remotely. Students can call 847-491-2151 to speak with the counselor on call.

Contact Information: In most situations, please contact us through Canvas (private messaging works well). If for some reason this is not possible, please feel free to email us via:

- hjmshi@u.northwestern.edu (Instructor)
- ShenyinyingTu2021@u.northwestern.edu (TA)

We are looking forward to a great Spring quarter!