Northeastern University
College of Engineering

Multidisciplinary Graduate Engineering
Course Syllabus

Course Information
INFO 6105-Data Science engineering Methods and Tools

Instructor Information
Dino (Constantin) Konstantopoulos
dino.k@northeastern.edu
Office Hours TBA

Textbook

Mathematics for Machine Learning, ISBN: 110845514X, ISBN13: 9781108455145, by Marc Peter Deisenroth (Author), A. Aldo Faisal (Author), Cheng Soon Ong (Author)

Course Prerequisites
Please consult school listings for prerequisites

Course Description
You will learn the math foundations and basic tools of Data Science while mastering Python. You will learn how to use multi-dimensional arrays and think in vectors, matrices, and tensors. You will learn how to operate on time series, spreadsheets, and higher dimensions. You will learn basic theories in probability, Bayesian statistics, linear algebra, and Machine learning (ML) by leveraging the 4 basic Data Science libraries written for Python: NumPy, Pandas, SciPy, and Scikit-learn. This class gives you the fundamental knowledge for applying for jobs that involve data analysis, such as jobs in the life sciences, financial, advertising, and social Web industries, and to be able to advance to specialized topics in Machine Learning.

Numpy adds Python support for large multi-dimensional arrays and matrices, along with a library of high-level mathematical functions to operate on these arrays. It focuses on fast number calculations, reads in fixed datatypes, improves RAM efficiency, and teaches you to think in Vectors. Pandas adds support for more refined data manipulation and analysis. It adds support for data structures and operations for manipulating tables and time series. SciPy is a collection of classic math and science algorithms and helper functions built on top of Numpy, such as linear and nonlinear regression, numerical optimization, etc. If you know the rules for dealing with your data, SciPy is the library for you. If you want the computer to learn the rules instead, and give you probabilistic answers, then

Scikit-learn, built on top of SciPy, is what you need. It is a Python module for machine learning at a basic level. If you can solve a problem with the methods in SciPy, it's more straightforward. If you can't, there's a good chance your problem is solvable using methods from Scikit-learn. We finish with an introduction to famous Machine Learning frameworks like TensorFlow, Torch, and Keras, and use cases like Natural Language Understanding and Vision.

The basic languages of Data Science are R and Python. We will start with a one-class introduction of R to get used to manipulating spreadsheets instead of single-entity variables, and then we switch to Python for the rest of the semester. A programming background in one managed language (C#, Java, or Python) is required, otherwise this class will be extremely hard for non-programmers. Your knowledge of Python will improve to black-belt level.

Grade is based on homework (30%), midterm (30%), a final project (30%), and a final exam (10%).

Student Learning/Course Outcomes (SLOs)

This class gives you the fundamental knowledge for applying for jobs that involve data analysis, such as jobs in the life sciences, financial, advertising, and social Web industries, and to be able to advance to specialized topics in Machine Learning (ML).

Attendance Policy

Students are expected to complete course readings, participate in class discussions or other learning activities during the unit, and complete written assignments for each unit during the time of that unit.

Late Work Policy

Students must submit assignments by the deadline in the time zone noted in the syllabus.
Students must communicate with TAs and faculty prior to the deadline if they anticipate work will be submitted late. Work submitted late without prior communication will loose points in proportion to how late it is submitted.

Grading/Evaluation Standards

Grade Breakdown:

30% Homework, 30% Midterm, 30% Final Project, 10% Final Exam.

Course Schedule

Homework is due next week of assignment, *before* class.

Academic Integrity

A commitment to the principles of academic integrity is essential to the mission of Northeastern University. The promotion of independent and original scholarship ensures that students derive the most from their educational experience and their pursuit of knowledge. Academic dishonesty violates the most fundamental values of an intellectual community and undermines the achievements of the entire University.

As members of the academic community, students must become familiar with their rights and responsibilities. In each course, they are responsible for knowing the requirements and restrictions regarding research and writing, examinations of whatever kind, collaborative work, the use of study aids, the appropriateness of assistance, and other issues. Students are responsible for learning the conventions of documentation and acknowledgment of sources in their fields. Northeastern University expects students to complete all examinations, tests, papers, creative projects, and assignments of any kind according to the highest ethical standards, as set forth either explicitly or implicitly in this Code or by the direction of instructors.

Go to http://www.northeastern.edu/osccr/academic-integrity-policy/ to access the full academic integrity policy.

Student Accommodations

Northeastern University and the Disability Resource Center (DRC) are committed to providing disability services that enable students who qualify under Section 504 of the Rehabilitation Act and the Americans with Disabilities Act Amendments Act (ADAAA) to participate fully in the activities of the university. To receive accommodations through the DRC, students must provide appropriate documentation that demonstrates a current substantially limiting disability.

For more information, visit http://www.northeastern.edu/drc/getting-started-with-the-drc/.

Library Services

The Northeastern University Library is at the hub of campus intellectual life. Resources include over 900,000 print volumes, 206,500 e-books, and 70,225 electronic journals.

For more information and for Education specific resources, visit http://subjectguides.lib.neu.edu/edresearch.

Diversity and Inclusion

Northeastern University is committed to equal opportunity, affirmative action, diversity and social justice while building a climate of inclusion on and beyond campus. In the classroom, member of the University community work to cultivate an inclusive environment that denounces discrimination through innovation, collaboration and an awareness of global perspectives on social justice.

Please visit http://www.northeastern.edu/oidi/ for complete information on Diversity and Inclusion

TITLE IX

Title IX of the Education Amendments of 1972 protects individuals from sex or gender-based discrimination, including discrimination based on gender-identity, in educational programs and activities that receive federal financial assistance.

Northeastern's Title IX Policy prohibits Prohibited Offenses, which are defined as sexual harassment, sexual assault, relationship or domestic violence, and stalking. The Title IX Policy applies to the entire community, including male, female, transgender students, faculty and staff.

In case of an emergency, please call 911.
Please visit www.northeastern.edu/titleix for a complete list of reporting options and resources both on- and off-campus.