MSML 605: Computing Systems for Machine Learning

Spring 2020

Instructor:

Mohammad Nayeem Teli, nayeem@cs.umd.edu.
Office: IRB 1128. Office phone: 301-405-0405.

Course Overview: This course will deal with the programming, software and hardware design and implementation issues of computing systems for machine learning. Topics in the programming/software domain will include: basic Python program structure, variables and assignment, built-in data types, flow control, functions and modules; basic I/O, and file operations. Classes, object-oriented programming and exceptions. Regular expressions, database access, network programming and sockets. Introduction to the Numpy, Scipy and Matplotlib libraries. Topics in the hardware domain include computer architecture, CPUs, single- and multi-core architectures, GPUs, memory and I/O systems, persistent storage, and virtual memory. Parallel processing architectures, multiprocessing and cluster processing.

Text: There is no recommended text book for the course. Most of the topics will be covered in the class from multiple sources and references will be provided whenever needed.

Requirements: Students enrolled in the course should be comfortable with programming (for those at UMD, having passed CMSC216 will be good enough!) and be reasonably mathematically mature. The course itself will make heavy use of the Python scripting language by way of Jupyter Notebooks, leaning on the Anaconda package manager; we'll give Python lectures, so don't worry if you haven't used Python before. Later lectures will delve into libraries necessary for statistics and machine learning and may make use of basic calculus and basic linear algebra; light mathematical maturity is preferred at roughly the level of a junior CS student.

Course Work and Exams: Course work will consist of written and programming homework assignments, one midterm exam, and a final. You may discuss homework problems and general solution strategies with classmates, but you ***must*** write up the solutions yourself.

Homework assignments will be turned in on ELMS, https://elms.umd.edu/.

As a courtesy to the grader, homeworks are to be written clearly and neatly and the programming assignments should be commented properly. Poorly written work will not be graded.

Exams:

The midterm exams will be:

Thursday, April 9, in Lecture

The final exam will be:

Thursday, May 14, in Lecture

Piazza: We will be using Piazza (www.piazza.com), a question-and-answer system designed to streamline discussion outside of the classroom. It supports LaTeX, code formatting, embedding of images, and attaching of files. It will be moderated by the instructors and TAs, but students are encouraged to answer questions.

ELMS: We will be using ELMS to submit assignments, and, post grades and to see recorded lectures (if any).

Grading: Final grades will be based on the written and programming assignments, quizzes, midterm exam, and a final exam. The weights of these will be approximately 1% for each quiz, 30% for assignments, 25% for midterm, and the remaining percentage (about 30-35%) for the final exam.

Late Penalty: You may submit your homework / programming assignment late by one day with a penalty of 20% off of your actual grade. However, there is one, get of jail free card. That means you may submit only one homework / programming assignment late by a day without any penalty.

Disability Support Services: Any student eligible for and requesting reasonable academic accommodations due to a disability is requested to provide, to the instructor in office hours, a letter of accommodation from the Office of Disability Support Services (DSS) within the first two weeks of the semester.

Course Evaluations: The Department of Computer Science takes the student course evaluations very seriously. Evaluations will usually be open during the last few weeks of the course. Students can go to www.courseevalum.umd.edu to complete their evaluations.

Syllabus: This is the current version of the syllabus. The instructors reserve the right to change it at any time.

Topics: The following is a *tentative* list of topics and readings in *approximate* order.

1. Introduction,
2. Python Porgramming,
3. Python data structures,
4. Numpy, Scipy,
5. Matplotlib,
6. Pandas,
7. Stats Library,
8. Github,
9. Database connectivity,
10. Tensorflow,
11. Hardware for ML.