

CMSC 320 - Introduction to Data Science

Syllabus

Instructor

Mohammad Nayeem Teli, nayeem@umd.edu, Section 0201.
Office: IRB 2224

Description

Welcome to CMSC 320. Data science encapsulates the interdisciplinary activities required to create data-centric products and applications that address specific scientific, socio-political or business questions. It has drawn tremendous attention from both academia and industry and is making deep inroads in industry, government, health and journalism – just ask Nate Silver!

This course focuses on data science pipeline, (i) data collection, (ii) exploratory and statistical data analysis, (iii) Machine Learning approaches, (iv) data and information visualization, and (v) the presentation and communication of analysis results. It will be centered around case studies drawing extensively from applications, and will yield a publicly-available final project that will strengthen course participants' data science portfolios.

This course will consist primarily of sets of self-contained lectures and assignments that leverage real-world data science platforms when needed; as such, there is no assigned textbook. Each lecture will come with links to required reading, which should be done before that lecture, and (when appropriate) a list of links to other resources on the web.

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 some Python-for-data-science primer lectures early on. Later lectures will delve into 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.

There will be two written, in class midterm examinations during lecture. In the interest of building students' public portfolios, and in the spirit of "learning by doing", students will create a self-contained online tutorial to be posted publicly on Github. This tutorial can be created individually or in a small group (at most 4). The tutorial will be a publicly-accessible website that provides an end-to-end walkthrough of identifying and scraping a specific data source, performing some exploratory analysis, and providing some sort of managerial or operational insight from that data.

Grading

Final grades will be computed according to the following weights. (These weights are **tentative** and subject to future adjustment.)

Percentage	Component
10%	weekly quizzes
20%	Midterm I
20%	Midterm II
35%	mini-project assignments (2% for Proj 0 and rest equally distributed between other projects)
15%	final tutorial to be posted publicly online

Late Penalty

Late assignments are not allowed.

Academic Honesty

Note that academic dishonesty includes not only cheating, fabrication, and plagiarism, but also includes helping other students commit acts of academic dishonesty by allowing them to obtain copies of your work. You are allowed to use the Web for reference purposes, but you may not copy code from any website or any other source. In short, all submitted work must be your own.

Cases of academic dishonesty will be pursued to the fullest extent possible as stipulated by the [Office of Student Conduct](#). Without exception every case of suspected academic dishonesty will be referred to the Office. If the student is found to be responsible of academic dishonesty, the typical sanction results in a special grade "XF", indicating that the course was failed due to academic dishonesty. More serious instances can result in expulsion from the university. If you have any doubt as to whether an act of yours might constitute academic dishonesty, please contact your TA or the course coordinator.

The University of Maryland, College Park has a nationally recognized Code of Academic Integrity, administered by the Student Honor Council. This code sets standards for academic integrity at Maryland for all undergraduate and graduate students. As a student you are responsible for upholding these standards for this course. It is very important for you to be aware of the consequences of cheating, fabrication, facilitation, and plagiarism. For more information on the Code of Academic Integrity or the Student Honor Council, please visit <http://www.shc.umd.edu>.

Examples of Academic Integrity Violations

The following are examples of academic integrity violations:

- Hardcoding of results in a project assignment. Hardcoding refers to attempting to make a program appear as if it works correctly (e.g., printing expected results for a test).
- Using any code available on the internet/web or any other source.
- Hiring any online service to complete an assignment for you.
- **You may not post the implementation of your assignments, materials related to the class (e.g., project description), or any other material associated with this online course except the final project. Even if the class is over and you have graduated, you may NOT post any material other than what is allowed.**
- Sharing your homework solutions or your "test code" with any student.
- Providing ideas/suggestions on how to solve/implement an assignment.
- Looking at or debugging another student's code.
- Using online forums to ask for help regarding our assignments.

Class Announcements

You are responsible for reading the class announcements that are posted on the webpage, Piazza and Canvas. Please check them often (at least once a day). Important information about the course (e.g., deadlines, assignment updates, etc.) will be posted on these platforms.

Excused Absence and Academic Accommodations

1. Any student who needs to be excused for an absence from a single class session, due to a medically necessitated absence shall:
 - **Make a reasonable attempt to inform the instructor of his/her illness prior to the class.** If you are going to miss an in-class assignment / quiz then we expect to hear from you (either email or telephone message) before the class session begins.
 - Upon returning to the class, present their instructor with a self-signed note attesting to the date of their illness. The note must contain an acknowledgment by the student that the information provided is true and correct. Providing false information to University officials is prohibited under Part 9(h) of the Code of Student Conduct (V-1.00(B) University of Maryland Code of Student Conduct) and may result in disciplinary action.
 - **This self-documentation may not be used for the Major Scheduled Grading Events as defined below.**
2. Any student who needs to be excused for more than one absence, or for a "Major Scheduled Grading Event", must provide written documentation of the illness from the Health Center or from an outside health care provider. This documentation must verify dates of treatment and indicate the timeframe that the student was unable to meet academic responsibilities. The documentation should be given to the instructor, not the TA. **We will not accept a "self-signed" note for "major scheduled grading events", as defined below. The note must be signed by a health care professional.**

The Major Scheduled Grading Events for this course include:

- Midterms
- Quizzes
- Homeworks
- Final project

It is also the student's responsibility to inform the instructor of any intended absences from exams for religious observances **in advance**. Notice should be provided as soon as possible but no later than one week prior to the exam.

Accessibility and Disability Support

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.

Copyright

All course materials are copyright UMCP, Department of Computer Science © 2024. All rights reserved. Students are permitted to use course materials for their own personal use only. Course materials may not be distributed publicly or provided to others (excepting other students in the course), in any way or format.