Berkeley
UNIVERSITY OF CALIFORNIA

Master of Information and Data Science (MIDS)
DATA SCI W255 Machine Learning Systems Engineering

Name	Role	Email	Slack
James York-Winegar	Course lead	winegarj@berkeley.edu	jameswine

Table of Contents:

Course Description

This course provides learners hands-on data management and systems engineering experience using containers, cloud, and Kubernetes ecosystems based on current industry practice. The course will be project based with an emphasis on how production systems are used at leading technology-focused companies and organizations. During the course, learners will build a body of knowledge on data management, architectural design, developing batch and streaming data

pipelines, scheduling, and data security, including access management and auditability. We'll also cover how these tools are changing the technology landscape.

Course Goals and Objectives

By the completion of this course, students will be able to:

- Describe the difference between a monolithic and microservice architecture, and assess and select appropriate use cases for each.
- Understand the continuous integration and continuous delivery (CI/CD) pipeline for automated code deployment, particularly for ML models.
- Describe the differences between a development and a production system, particularly for machine learning where the boundaries are blurry.
- Understand how to serve machine learning models over an API in real time.
- Know when to leverage a cache for serving machine learning models to reduce load on production systems.
- Understand how stateful systems add complexities to systems engineering.
- Demonstrate understanding of Kubernetes for management of machine learning models.
- Construct, measure, and identify metrics relating to performance of a system in order to optimize costs and latency of serving inferences for machine learning models.

Key Dates

Academic calendar: https://www.ischool.berkeley.edu/intranet/students/mids/calendar

Course Evaluation

Assignment	Grading Scale	Weight
Lab 1	100 points	10%
Lab 2	100 points	10%
Lab 3	100 points	10%
Lab 4	100 points	10%
Lab 5	100 points	10%
Final Project	100 points	50%

All labs and projects will be published to I School GitHub for W255: https://github.com/UC-Berkeley-I-School/mids-w255.

All labs will be due on Monday 11:59 p.m. PST of the corresponding week indicated due.

General Grading Philosophy

The course will be graded on an absolute scale, and the grades will not be fitted to a specific curve. This is a graduate-level course, and we trust that different students will have varying levels of interests in the different subjects in the course. As such, the grading scheme is designed to acknowledge this intellectual diversity.

Late Submission Policy

Solutions of labs will be discussed during the live sessions of the course. Therefore, any assignment that is submitted after the deadline will be returned without grading and will receive a grade of zero.

Collaboration Policy/Academic Integrity

All students—undergraduate, graduate, professional full time, part time, law, etc.—must be familiar with and abide by the provisions of the "Student Code of Conduct," including those provisions relating to Academic Misconduct. All forms of academic misconduct, including cheating, fabrication, plagiarism, or facilitating academic dishonesty, will not be tolerated. The full text of the UC Berkeley Honor Code is available at https://teaching.berkeley.edu/berkeley-honor-code, and the Student Code of Conduct is available at https://sa.berkeley.edu/student-code-of-conduct#102.01_Academic_Misconduct.

*We encourage studying in groups of two to four people. This applies to working on homework, discussing labs and projects, and studying for the exam. However, students must always adhere to the UC Berkeley Code of Conduct (http://sa.berkeley.edu/code-of-conduct) and the UC Berkeley Honor Code (https://teaching.berkeley.edu/berkeley-honor-code). In particular, **all materials that are turned in for credit or evaluation must be written solely by the submitting student or group**. Similarly, you may consult books, publications, or online resources to help you study. In the end, **you must always credit and acknowledge all consulted sources in your submission (including other persons, books, resources, etc.).***

Weekly Schedule Outline/Assignments
- Week 1: Monolith and Microservice Architectures
- Week 2: Continuous Integration and Continuous Delivery
- Week 3: Development and Production Systems
- Week 4: Machine Learning Models Over an API, Part I
  - Lab 1 due, Monday 11:59 p.m. PST
- Week 5: Machine Learning Models Over an API, Part II
- Week 6: Cache

- - Lab 2 due, Monday 11:59 p.m. PST
  - Week 7: State
  - Week 8: Kubernetes, Part I
  - Week 9: Kubernetes, Part II
  - Week 10: Kubernetes, Part III
    - Lab 3 due, Monday 11:59 p.m. PST
  - Week 11: Deploy and Instrument a Machine Learning Prediction API
    - Lab 4 due, Monday 11:59 p.m. PST
  - Week 12: Model eEvaluation
    - Lab 5 due, Monday 11:59 p.m. PST
  - Week 13: Security
    - Final Project Presentation due 24 hours prior to the final live session

Attendance and Participation

We believe in the importance of the social aspects of learning—between students, and between students and instructors—and we recognize that knowledge-building is not solely occurring on an individual level but that it is built by social activity involving people and by members engaged in the activity. Participation and communication are key aspects of this course that are vital to the learning experiences of you and your classmates.

Therefore, we like to remind all students of the following requirements for live class sessions:

- You are required to join live class sessions from a study environment with video turned on and with a headset for clear audio, without background movement or background noise, and with an internet connection suitable for video streaming.

- You are expected to engage in class discussions, breakout room discussions, and exercises, and to be present and attentive for your and other teams' in-class presentations.

- Keep your microphone on mute when not talking to avoid background noise. Do your best to minimize distractions in the background video, and ensure that your camera is on while you are engaged in discussions.

That said, in exceptional circumstances, if you are unable to meet in a space with no background movement, or if your connection is poor, make arrangements with your instructor (beforehand if possible) to explain your situation. Sometimes connections and circumstances make turning off video the best option. If this is a recurring issue in your study environment, you are responsible for finding a different environment that will allow you to fully participate in classes, without distraction to your classmates. Please contact Student Affairs if you have problems meeting these requirements.

Attendance at all live class meetings is mandatory. All students should arrive a few minutes before class ready to participate when class starts. Preparation for class is mandatory. All students should have watched all of the asynchronous videos, read all the required readings, worked through all of the examples and proofs in the readings, and worked through all of the Jupyter Notebooks assigned prior to class each week. Active class participation is mandatory, which can only be achieved when students come to class prepared.

Students should ensure they have all the technical requirements in terms of a high-speed, high-quality internet connection, an area free from background noise, audio equipment that cancels both background noise and suppresses any echoes, etc. Students should ensure that they know how to share their desktop by practicing this prior to the first class meeting.

Failure to adhere to these requirements will result in an initial warning from your instructor(s), followed by a possible reduction in grades or a failing grade in the course.

Diversity and Inclusion

Integrating a diverse set of experiences is important for a more comprehensive understanding of cybersecurity. We will make an effort to read papers and hear from a diverse group of practitioners. Still, limits exist on this diversity in the field of data science. I acknowledge that it is possible that there may be both overt and covert biases in the material due to the lens with which it was created. I would like to nurture a learning environment that supports a diversity of thoughts, perspectives, and experiences, and honors your identities (including race, gender, class, sexuality, religion, ability, veteran status, etc.) in the spirit of the UC Berkeley Principles of Community: https://diversity.berkeley.edu/principles-community.

To help accomplish this, please contact me or submit anonymous feedback through I School channels if you have any suggestions to improve the quality of the course. If something was said in class (by anyone) or you experience anything that makes you feel uncomfortable, please talk to me about it. If you feel like your performance in the class is being impacted by experiences outside of class, please don't hesitate to come and talk with me. I want to be a resource for you. Also, anonymous feedback is always an option, and may lead to me to make a general announcement to the class, if necessary, to address your concerns. As a participant in teamwork and course discussions, you should also strive to honor the diversity of your classmates.

If you prefer to speak with someone outside of the course, the MIDS Academic Director Drew Paulin, the I School Assistant Dean of Academic Programs Catherine Cronquist Browning, and the UC Berkeley Office for Graduate Diversity are excellent resources. Also see the following: https://www.ischool.berkeley.edu/about/community.

Disability Services and Accommodations

If you need disability-related accommodations in this class, if you have emergency medical information you wish to share with me, or if you need special arrangements in case the building must be evacuated, please inform me as soon as possible.

The I School recognizes disability in the context of diversity, and the Disabled Students' Program (DSP) equips students with appropriate accommodations and services to remove barriers to educational access. Students seeking accommodations in this class are responsible for completing the DSP application process to obtain an accommodation letter. (510) 642-0518, https://dsp.berkeley.edu.

Publishing Your Work

You are highly encouraged to use your program coursework to build an academic/professional portfolio.

- Blog about your coursework (and other ideas) and share on the I School Medium.
  - Instructions are here on the intranet for students: https://www.ischool.berkeley.edu/intranet/connect.
  - And here public for alumni: https://www.ischool.berkeley.edu/alumni/stay-connected.
- Publish projects to your I School project portfolio gallery (more than just for capstone).
- Publish your work on LinkedIn and tag the @UC Berkeley School of Information.
- Publish in academic journals—contact your professors for assistance. (Note that multiple review iterations are usually required; this can be a time-intensive endeavor.)
  - For help writing professional academic papers, students are encouraged to contact Sabrina Soracco, the Director of the Graduate Writing Center, in the Graduate Division—see https://grad.berkeley.edu/staff/sabrina-soracco/.
  - And see https://grad.berkeley.edu/professional-development/graduate-writing-center/ for links to resource guides, appointments with consultants, workshops, etc.
- Publish your news (e.g., conference talks, award, scholarships) to the I School internal newsletter.