Syllabus

Course Info

- **Instructor**: Bryan Wilder
- **TA**: Ananya Joshi
- **Meetings**: Tuesdays and Thursdays, 9:30-10:50am
- **Course website:** https://github.com/bwilder0/mlpractice_s2023

1. Course Description

This is a project-based course designed to provide students training and experience in solving real-world problems using machine learning, exploring the interface between research and practice. The goal of this course is to give students exposure to the nuance of applying machine learning to the real-world, where common assumptions (like iid and stationarity) break down. Students will learn how to formulate real-world business or policy scenarios as machine learning problems, how to address common challenges which arise in applying ML to such problems (e.g., distribution shift or missingness), and how to rigorously evaluate the results of such interventions in practice (e.g., through designing randomized trials or observational studies). We will place an emphasis throughout on issues related to ethics and fairness in machine learning, and discuss how choices throughout the machine learning pipeline – including problem formulation, outcome definition, data collection, and model training – contribute to the social impact of algorithmic systems.

For more details about topics covered, see the Course Schedule on the website.

2. Prerequisites

Students entering the class are expected to have a pre-existing working knowledge of common machine learning methods, as well as implementation in python. We will focus on how to choose between different machine learning techniques, modify them to address common real-world challenges, and evaluate their performance. We will assume that students are already familiar with the models themselves (e.g., linear regression, decisions trees, ensembles, neural networks), as well as how to train such models in python.

3. Recommended Textbooks

The core content of this course does not exactly follow any one textbook. We will have required and suggested readings associated with each lecture, which can be found on the course website. Two general references that may be helpful:

- AI for Social Impact. Edited by Fang, Tambe, and Wilder. Freely available at https://ai4sibook.org/
- Big Data and Social Science. Edited by Foster, Ghani, Jarmin, Kreuter and Lane. Freely available at https://textbook.coleridgeinitiative.org/

4. Course Components

Grading

The requirements of this course consist of participating in lectures, midterm and final exams, homework assignments, and readings. The grading breakdown is the following:

- Weekly project update assignments (25%)
- Midterm take-home exam (20%)
- Write-up on distribution shift findings (15%)
- Write-up on fairness findings (15%)
- Final project reflection write-up (10%)
- Paper reflections (10%)
- Class attendance and participation (5%)

Midterm exam

The course will have a take-home midterm exam. You may use any resources you would like (course readings, slides, books, the internet, etc) during the course of solving the exam. However, you may not communicate with any other person, including classmates, about the exam in any way. The work must be entirely your own.

Participation

This course emphasizes the process of thinking through real-world problems and how and when they can be addressed using machine learning. Accordingly, our class sessions will rely on student participation to discuss potential scenarios and case studies together. You are highly encouraged to engage actively during class discussions, which will make the course much more exciting for everyone. Participation will also factor into your grade as noted above. In order to prepare, you are expected to have read all required readings before each class. During such discussions, we expect you to be respectful at all times towards your fellow students.

Attendance

You are expected to attend class in order to participate (as noted above). If you cannot attend on a given day, email the instructor in advance. Examples of acceptable reasons are attending a conference, having an infectious disease, or family or medical emergencies (discussed further below under “Extensions”).

5. Technologies

We use a variety of technologies:

Slack

We will use Slack for all **course discussion**. Questions about homeworks, course content, logistics, etc. should all be directed to Slack. Please use the #questions channel to post questions. If you have a question, chances are several others had the same question. By posting your question publicly on Slack, the course staff can answer once and everyone benefits. If you have a private question, you should also use Slack as it will likely receive a faster response.

Canvas

We use Canvas to collect submissions of reports, paper reflections, and the take-home midterm.

Regrade Requests: If you believe an error was made during grading, please send an email to the TA. For each assignment, regrade requests will be open for only 1 week after the grades have been published. This is to encourage you to check the feedback you’ve received early!

6. General Policies

Late submissions

Each project group receives a total of 4 grace days that can be distributed among any of the weekly project updates or the fairness or distribution shift reports. We will automatically apply these late days when an assignment is not received. Once all grace days have been used, further late submissions will receive a zero.

No late days can be used for the midterm or for the final reflection; late submissions will receive a zero. No late days can be used for the paper reflections (late submissions will receive a zero here as well). However, we will drop the lowest grade among the paper reflections when calculating the average for that portion of the grade.

Extensions

In general, we do not grant extensions on assignments. There are several exceptions:

- **Medical Emergencies:** If you are sick and unable to complete an assignment or attend class, please go to University Health Services. For minor illnesses, we expect grace days or our late penalties to provide sufficient accommodation. For medical emergencies (e.g. prolonged hospitalization), students may request an extension afterwards and should include a note from University Health Services.
- **Family/Personal Emergencies:** If you have a family emergency (e.g. death in the family) or a personal emergency (e.g. mental health crisis), please contact your academic adviser or Counseling and Psychological Services (CaPS). In addition to offering support, they will reach out to the instructors for all your courses on your behalf to request an extension.
- **University-Approved Absences:** If you are attending an out-of-town university approved event (e.g. multi-day athletic/academic trip organized by the university), you may request an extension for the duration of the trip. You must provide confirmation of your attendance, usually from a faculty or staff organizer of the event.

For any of the above situations, you may request an extension **by emailing your instructor**. The email should be sent as soon as you are aware of the conflict and at least **5 days prior to the deadline**. In the case of an emergency, no notice is needed.

Audit Policy

Official auditing of the course (i.e. taking the course for an “Audit” grade) is not permitted this semester.

Unofficial auditing of the course (i.e. watching the lectures online or attending them in person) is welcome and permitted without prior approval. We give priority to students taking the course for a letter grade, so auditors may only take a seat in the classroom if there is one available *10 minutes after the start of class*. Unofficial auditors will not be given access to course materials such as homework assignments and exams.

Pass/Fail Policy

We allow you to take the course as Pass/Fail. Instructor permission is not required. What grade is the cutoff for Pass will depend on your program. Be sure to check with your program / department as to whether you can count a Pass/Fail course towards your degree requirements.

Accommodations for Students with Disabilities:

If you have a disability and have an accommodations letter from the Disability Resources office, I encourage you to discuss your accommodations and needs with me as early in the

semester as possible. I will work with you to ensure that accommodations are provided as appropriate. If you suspect that you may have a disability and would benefit from accommodations but are not yet registered with the Office of Disability Resources, I encourage you to contact them at access@andrew.cmu.edu.

7. Academic Integrity Policies

Read this carefully!

Collaboration among Students

- **Group project**: students are expected to collaborate with their groupmates on the project and implementation. Students may discuss the project with their classmates in other groups, provided that no written materials (course reports, code, etc) are shared. The implementation of the project and the course deliverables should be entirely the work of the group members.
- **Paper reflections:** Students may discuss readings with their classmates provided that no written materials (e.g., drafts of the assignment) are shared. The final written product should be entirely the work of each individual.
- **Midterm exam:** No collaboration or discussion is permitted with any other person.
- Collaboration outside these guidelines will be handled severely, in compliance with CMU's Policy on Academic Integrity.

Previously Used Assignments

Some of the homework assignments used in this class may have been used in prior versions of this class, or in classes at other institutions, or elsewhere. Solutions to them may be, or may have been, available online, or from other people or sources. It is explicitly forbidden to use any such sources, or to consult people who have solved these problems before. It is explicitly forbidden to search for these problems or their solutions on the internet. Your group must solve the project assignments completely on your own.

Policy Regarding "Found Code":

You are encouraged to read books and other instructional materials, both online and offline, to help you understand the concepts and algorithms taught in class. These materials may contain example code or pseudo code, which may help you better understand an algorithm or an implementation detail. However, when you implement your own solution to an assignment, you must put all materials aside, and write your code completely on your own, starting "from scratch". Specifically, you may not use any code you found or came across. If you find or come across code that implements any part of your assignment, you must disclose this fact in your collaboration statement.

Duty to Protect One's Work

Students are responsible for pro-actively protecting their work from copying and misuse by other students. If a student's work is copied by another student, the original author is also considered to be at fault and in gross violation of the course policies. It does not matter whether the author allowed the work to be copied or was merely negligent in preventing it from being copied. When overlapping work is submitted by different students, both students will be punished.

To protect future students, do not post your work publicly, neither during the course nor afterwards.

Penalties for Violations of Course Policies

All violations (even first one) of course policies will always be reported to the university authorities (your Department Head, Associate Dean, Dean of Student Affairs, etc.) as an official Academic Integrity Violation and will carry severe penalties.

1. The penalty for the first violation is a one-and-a-half letter grade reduction. For example, if your final letter grade for the course was to be an A-, it would become a C+.
2. The penalty for the second violation is failure in the course, and can even lead to dismissal from the university.

8. Support

Take care of yourself. Do your best to maintain a healthy lifestyle this semester by eating well, exercising, avoiding drugs and alcohol, getting enough sleep and taking some time to relax. This will help you achieve your goals and cope with stress.

All of us benefit from support during times of struggle. You are not alone. There are many helpful resources available on campus and an important part of the college experience is learning how to ask for help. Asking for support sooner rather than later is often helpful.

If you or anyone you know experiences any academic stress, difficult life events, or feelings like anxiety or depression, we strongly encourage you to seek support. Counseling and Psychological Services (CaPS) is here to help: call 412-268-2922 and visit their website at http://www.cmu.edu/counseling/. Consider reaching out to a friend, faculty or family member you trust for help getting connected to the support that can help.

If you or someone you know is feeling suicidal or in danger of self-harm, call someone immediately, day or night:

- CaPS: 412-268-2922

- Re:solve Crisis Network: 888-796-8226
- If the situation is life threatening, call the police:
    - On campus: CMU Police: 412-268-2323
    - Off campus: 911.

If you have questions about this or your coursework, please let the instructors know.

---

9. Note to people outside CMU

Please feel free to reuse any of these course materials that you find of use in your own courses. We ask that you retain any copyright notices, and include written notice indicating the source of any materials you use.