EECS 553 WINTER 2026 SYLLABUS DRAFT, MACHINE LEARNING - ECE*

The goal of machine learning is to develop computer algorithms that can learn from data or past experience to make accurate, useful, and fair predictions on new unseen data. The capabilities and successes of machine learning have grown tremendously over the last few decades, and extremely rapidly over the past few years. This has had major impacts in many real-world applications, yet other applications have seen little to no progress. This course will give a graduate-level introduction of machine learning and provide mathematical and statistical foundations of machine learning, mathematical derivation and implementation of the algorithms, and discussion of their applications. On assignments, students will work through mathematical derivations applying principles learned in class, implement machine learning algorithms in Python, and apply those algorithms to data sets spanning a variety of applications. In small groups, students will also participate in a machine learning reproducibility challenge, which involves reproducing the results and assessing the conclusions of a paper published in a major machine learning conference.

Instructor: Professor Laura Balzano, girasole@umich.edu

Course time: Tuesdays and Thursdays, 12-1:30pm US Eastern time

Course location: TBD. All lectures will be recorded.

GSI(s): TBD

Advisory Pre-requisites: Graduate level coursework in probability and linear algebra. A curiosity about the mathematical models and representations behind machine learning algorithms. Patience with abstraction and with the required effort to translate mathematical models into algorithms and applications.

Required Textbook: None. All homeworks and quizzes will be based on lecture notes and slides.
Additional useful texts are listed on Canvas>Files>MLTexts.pdf
Office hours: TBD

Communication and Course Admin: The course will be run through Canvas. All materials will be posted there. Announcements in Canvas will be the main venue for communication from the teaching staff to students. Lecture recordings will be available through Canvas. Less critical messages from teaching staff may also be sent through Piazza. Piazza through Canvas will be used for class discussions. Gradescope will be linked there for homework submission. iClicker will be used for in-class participation.

Grade breakdown:
Exams 40%, HW 15%, ML Reproducibility Project 30%, Floater 5% takes on your highest grade in the above three categories
Clicker questions 5%, Piazza participation 5%

Exams: There will be two in-person exams.
You will be allowed two note sheets (8.5×11", both sides), but it is otherwise closed-book. The exam will be mostly multiple choice. You will have access to a practice exam one week before the exam.

If you have other mandatory obligations that conflict with the exam date, please inform the entire teaching staff with a private Piazza message using subject "EECS 553 - exam conflict" as soon as possible and two weeks before the exam at the latest. Those students will take the exam during class time on the same day. If you have SSD accommodations for exams, please send your documentation to Professor Balzano with the subject line "EECS 553 - SSD" at the start of the semester, two weeks before the first exam at the latest. If you send the documentation through the "Accommodate" system, please also send the professor an email with the above subject line, which helps me quickly find you when the exam time comes.

*This is an tentative version as of November 21, 2025.

If you get sick before the exam, please inform the entire teaching staff with a private Piazza message as soon as possible with subject line "EECS 553 - illness". If it is within 3 hours of the start time, please send an email to Professor Balzano and the rest of the teaching staff with subject line "EECS 553 - illness". Either way you must send your message before the exam. We will quickly arrange an alternative option for you.

Homework: Homework policies and deadlines TBD.

Every homework will ask the students to submit at least one multiple choice question on the current subject matter along with its solution. These are graded only on whether both a question and answer were provided. These questions with solutions will be posted in Perusall for discussion, typo identification, etc. All proposed questions will be visible to all other students and could be used for the exams. Simple questions are not discouraged; more important is that the question be correctly posed with a correct answer provided. If the question is too vague, posed incorrectly, has no correct answer, or is otherwise faulty it may be removed from shared materials.

Project: In small groups, students will participate in a machine learning reproducibility challenge, which involves reproducing the results and assessing the conclusions of a paper published in a major machine learning conference. This will be a group project that you will work on throughout the semester with 3-4 students per team. I recommend you choose your group to have people with experience in both programming and theory. I opened the "search for teammates" Piazza post for you to find groups.

Deadlines for project assignments (PA) are as follows:
TBD
The ML community has had an annual Machine Learning reproducibility challenge, whose goal "is to encourage the publishing and sharing of scientific results that are reliable and reproducible. In support of this, the objective of this challenge is to investigate reproducibility of papers accepted for publication at top conferences by inviting members of the community at large to select a paper, and verify the empirical results and claims in the paper by reproducing the computational experiments, either via a new implementation or using code/data or other information provided by the authors." (Quote from `https://paperswithcode.com/rc2022`.)

For the project in this class, each small group of 3-4 students will take part in the challenge. **For class, you may select any paper to reproduce from NeurIPS, ICML, ICLR, AAAI, AI Stats, as well as the journals JMLR and TMLR from 2020-2026.** The paper should have a major machine learning component for you to reproduce and evaluate their claims. You should highlight the machine learning aspects in your initial justification when selecting these papers and note that your selection is subject to approval by the instructional staff.

"Reproducing" a paper for the challenge involves investigating their *claims and conclusions* and evaluating whether those claims and conclusions are reproducible. That is likely to include reproducing their experimental results on both synthetic and real data. Interaction with the original authors is encouraged (and early on in your project if you expect a response before the deadline). Groups may use code that authors share online or by email. Students may also choose to implement the algorithms from scratch, especially if the paper claims that the algorithms are simple to implement. If the full data and code are available for the paper online, this may be somewhat straightforward, but not always due to compatibility issues or missing pieces.

However, even when all the data/code are available and easy to run and reproduce exact plots in the papers, claims often are made more broadly, yet in the paper they may be only illustrated on a few datasets. Therefore, reproducing their claims could involve applying their algorithm with different hyperparameters, to new datasets, and/or to synthetic data with realistic variations from the original paper. If computational efficiency is a focus of the original paper, students may want to time the algorithms using their own computing resources to see if they align with the author's numbers. If the algorithm has several components, (additional) ablation studies may be carried out. Claims may also involve mathematical statements that

should be carefully checked, but simply saying "We checked the proof and it is correct" is not enough – one must look at parts of the proofs or derivations where steps are missing, flesh out those steps, include additional references for justification, etc.

Participation: There will be two main opportunities for participation: Piazza and iClicker. If you are regularly active and engaged using these formats, you will get a high score for participation.

Piazza is used as a message board – a place to ask and answer questions, share thoughts, discuss ideas, applications, and algorithms, etc. Students should post a question/comment, share a resource, or answer a question – and it must be visible to the whole class, though anonymously is fine – once per week on average to get full credit. I especially encourage the sharing/discussion of typos in the notes or other shared sources. Scores will be tallied monthly.

For iClicker, I will periodically ask questions during class to make sure everyone is following along. Each question is worth 3 points. Each question will (hopefully) have one best answer, but may have multiple answers that count as "right." You will get 2 points for attempting the question and 3 points for getting one of the right answers. The **ten days** of lowest scores will be dropped.

Attendance: Attendance is encouraged but not required. This is designed to be flexible so you are not tempted to come in when you are sick. Nevertheless, please don't enroll for this class if you will be remote for much of the semester or have other classes or obligations during our class time. I will never make accommodations for a student who has double-booked themselves, and I promise it's not worth it to try to juggle two courses/obligations in the same time slot. iClicker is worth 5% of your grade and requires attendance, and attendance is required at quizzes and presentation days (both with exceptions for sickness listed above).

Computing resources: Students may remotely connect to CAEN lab computers. These computers are intended for interactive use, as opposed to long-running, resource intensive jobs.
`https://caen.engin.umich.edu/connect/`

In addition, for the final project, each student will be granted a certain budget of computing time on the Great Lakes computing cluster, which includes both CPUs and GPUs. Use of the cluster is entirely optional. Students are advised that scheduling jobs, especially GPU jobs, requires careful setup and the 553 instructors cannot guarantee support for all student questions. In other words, the service is offered "as is." That said, students may ask questions of the class through Piazza as well as the ARC-TS staff. Because of the time needed to set up jobs (again, especially GPU jobs), and delays in responses to getting technical questions answered, students should weigh the pros and cons of using the cluster. Starting early is definitely recommended. Difficulties using the cluster will not be a valid excuse for incomplete project experiments.

As an alternative to Great Lakes, students may also consider Google Colaboratory, which gives each student access to a free albeit slow GPU.

Honor policy: I take academic integrity very seriously. The College of Engineering is a community in which personal responsibility, honesty, fairness, respect, and mutual trust are maintained. You are expected to practice the highest possible standards of academic integrity. Any deviation from this expectation will result in a minimum academic penalty of your failing the assignment, and will result in additional disciplinary measures. This includes, but is not limited to, cheating, using unauthorized material during exams, using or copying another student's work, copying solutions from internet materials, plagiarism, and any other form of academic misrepresentation or misconduct. In particular, any 553 materials that are found on coursehero, chegg, etc will be immediately reported to the honor council.

For a list of actions that constitute misconduct, and possible sanctions for those actions, please see the Code of Conduct at `http://www.engin.umich.edu/college/academics/bulletin/rules`. I will not hesitate to enforce these standards in 553.

Use of Generative AI: The use of Generative AI (GenAI) tools (e.g., ChatGPT, GPT-5, Github Copilot, DALL-E, etc.) is allowed and encouraged in certain contexts in Professor Balzano's EECS 553 class. This is an important tool – built on the machine learning concepts you will see in this very class – that you will be expected to know how to use in industry when you graduate. I encourage you to use it in support of your learning and understanding and to accelerate implementation of minor tasks that would take longer without it. I also caution you from using it at a time when it might shortcut your own learning – what is known in education as the "productive struggle" that leads to a deeper conceptual understanding. You may feel like you are using it only when you are otherwise stuck, but it's actually when you mentally push through that feeling of being stuck that you actually learn the concept.

GenAI is not allowed in the following contexts:

1. Piazza discussion board posts and responses, as these are intended to be your own personal questions, answers, and reflections.
2. On any homework problem that does not explicitly use or ask about GenAI.
3. In particular don't use GenAI to generate the new multiple choice problems for our assignments – these should be from your own imagination. They don't have to be complicated, but they should be yours.
4. On any exam – the exams will be administered in-class, closed book, without calculators or any other technology.
5. For participation assignments, such as exam reflection documents or seminar summaries.

GenAI is permitted in the following contexts:

1. To answer problems that involve GenAI on the homework. You must save and submit the transcript of your session.
2. To help you study for exams, including using Generative AI generate new practice problems for studying, or help with creating a notes sheet for the exam, etc. I suggest you carefully check that everything it generates is correct.
3. For project work that is unsubmitted – e.g., you may be interested in using Generative AI to brainstorm ideas, identify data sources, or get ideas for metrics to showcase your algorithms.
4. GenAI for the submitted project assignments is permitted with disclosure. For example, you may ask it to rewrite a paragraph, find bugs in your already written code, suggest an outline for a report, etc. If you choose to use GenAI in submitted project work, you will include all the details in a "Project GenAI Disclosure Report" (see the project submission assignments for details). Save the transcript of your sessions, and track changes as you edit them, so that it is possible to create a complete disclosure report when the time comes. Use of Generative AI on the project is NOT required.

In all cases where GenAI is permitted,

- You need to be open about your use and document your use.
- Ultimately, you take full responsibility for AI-generated materials: calculations should be correct, ideas should be attributed, and stated facts should be true.
- This course's assignments are designed to help you practice and understand material: over-use of AI may thus inhibit your learning and long-term retention of the material.

Generative AI often makes mistakes (from simple math mistakes, to programming bugs, to gibberish or overly generic text responses, to conceptual errors or hallucinated facts). If you choose to use these tools in the context when it is allowed, it is of course your responsibility to read the outputs critically and find these mistakes. If you choose to use these tools to study for an exam and you perform poorly due to misleading Generative AI mistakes or hallucinations, that is also your responsibility.

Failure to follow these guidelines or use of Generative AI in a manner that violates the UM CoE Honor code will be considered a violation of academic integrity. Generative AI tools should be used to enhance understanding and creativity, not replace your original thought or fundamental skills such as critical thinking, problem-solving, and writing. Just like any other kind of cheating where you don't do the work yourself (e.g., copying from a friend or a solution on Chegg), the rampant use of Generative AI to replace your own

work will only shortchange your own learning at the end of the day. Should you have any questions, doubts, or concerns about using Generative AI tools in your coursework, or if you have specific situations where you would like to use it but you aren't sure if it's allowed, please talk to me in person or using Piazza.

Commitment to equal opportunity: The Faculty of the COE are committed to a policy of equal opportunity for all persons and do not discriminate on the basis of race, color, national origin, age, marital status, sex, sexual orientation, gender identity, gender expression, disability, religion, height, weight, or veteran status. Please feel free to contact your instructors with any problem, concern, or suggestion. We ask that all students treat each other with respect.

Diversity, Equity, and Inclusion: I consider this classroom to be a place where you will be treated with respect, and I celebrate individuals of all ages, backgrounds, beliefs, ethnicities, genders, gender identities, gender expressions, national origins, religious affiliations, sexual orientations, ability, and other visible and nonvisible differences. All members of this class are expected to contribute to a respectful, welcoming and inclusive environment for every other member of the class. I am dedicated to helping each of you achieve all that you can in this class. I may, either in lecture or smaller interactions, accidentally use language that creates offense or discomfort. Should I do this, please contact me to help me understand and avoid making the same mistake again. If you are someone whose background is underrepresented in electrical engineering and computer science, I want to be your advocate.

Accommodations for Students with Disabilities: If you think you need an accommodation for a disability, please let you instructor know at your earliest convenience. Some aspects of this course may be modified to facilitate your participation and progress. As soon as you make us aware of your needs, we can work with the Services for Students with Disabilities (SSD) office to help us determine appropriate academic accommodations. SSD (734-763-3000; `http://ssd.umich.edu`) typically recommends accommodations through a Verified Individualized Services and Accommodations (VISA) form. Any information you provide is private and confidential and will be treated as such.

Student Mental Health and Wellbeing: University of Michigan is committed to advancing the mental health and wellbeing of its students, as am I. While I am not a mental health expert, I have personal experience with mental health struggles, and I support you in identifying these challenges and seeking help. If you or someone you know is feeling overwhelmed, depressed, and/or in need of support, services are available. For help, contact Counseling and Psychological Services (CAPS) at (734) 764-8312 and `https://caps.umich.edu/` during and after hours, on weekends and holidays, or through its counselors physically located in schools on both North and Central Campus. You may also consult University Health Service (UHS) at (734) 764-8320 and `https://www.uhs.umich.edu/mentalhealthsvcs` or `https://uhs.umich.edu/stressresources`,
or for alcohol or drug concerns, see `www.uhs.umich.edu/aodresources`. For a listing of other mental health resources available on and off campus, visit:
`https://caps.umich.edu/article/um-mental-health-resources`.

Auditing: To audit the course, students should enroll as a visitor. There are no required assignments for auditors/visitors. Your program coordinator can help you enroll as a visitor. Please print this page to pdf and provide to your program coordinator as proof of our approval to enroll as a visitor. See also the Rackham policy for Visiting: `https://rackham.umich.edu/academic-policies/section3/`

Course plan (subject to change):

Thurs Jan 8 - Intro, Probabilistic Setup for Classification, Risk/Empirical Risk,
HW 1 released
Tues Jan 13 - Unconstrained Optimization
Thurs Jan 15 - Bayes Classifier, LDA, HW 1 on probability/matrices due, HW 2 released
Tues Jan 20 - QDA, Naive Bayes
Thurs Jan 22 - Logistic Regression, Linear Regression
Tues Jan 27 - Empirical Risk Minimization, Regularization
Thurs Jan 29 - Newton's method, Separating Hyperplanes, HW 2 due, HW 3 released
Tues Feb 3 - Constrained optimization
Thurs Feb 5 - Kernel methods, Support Vector Machine
Tues Feb 10 - Classtime office hour: discussion of ML paper selections
Thurs Feb 12 - Kernel SVM, Kernel Nearest Neighbors, Kernel Ridge Regression
Fri Feb 13 - Selection of ML paper due
Tues Feb 17 - Review and Q&A
Thurs Feb 19 - Exam 1, HW 4 released
Tues Feb 25 - Deep feedforward networks
Thurs Feb 26 - Backpropagation, Convolutional Neural Networks
== Winter Break ==
Tues March 10 - Unsupervised learning introduction, Dimensionality reduction, PCA
Thurs March 12 - Nonlinear dimensionality reduction, Clustering, K-means
Sun March 15 - First draft ML report due
Tues March 17 - Gaussian Mixture Model, Expectation-Maximization, HW 4 due
Thurs March 19 - Ensemble Methods
Tues March 24 - Transformers and Attention, HW 5 released
Thurs March 26 - Matrix factorization, Matrix completion, Low-Rank Adaptation
Tues March 31 - Anomaly detection, Distribution shift
Thurs Apr 2 - Socially Responsible Machine Learning
Tues Apr 7 - Review and Q&A, HW 5 due
Thurs Apr 9 - Exam 2
Tues Apr 14 - TBD
Thurs Apr 16 - TBD
Tues Apr 21 - TBD
Wednesday, April 29 8-10am is our exam time.
This will probably be a poster session for our final projects.