Mathematical Foundations of Machine Learning

MATH 4990V/5990V, Fall 2026

Instructor	Jiahui Chen
Email	jiahuic@uark.edu
Office	SCEN 228
Office Hours	Mon. & Fri. 11:40 AM–1:00 PM; Wed. 3:30–4:45 PM
Meeting Time	MWF, 8:35 AM–9:25 AM
Classroom	CHPN 425
Semester Dates	August 17 – December 3, 2026 (University of Arkansas Fall 2026 academic calendar)
Course Repository	https://git.uark.edu/jiahuic/math-499v-599v-mfml
Course Page	https://jiahuic.github.io/teaching/math-4990v5990v

Communication

Email is the best way to contact the instructor outside office hours. Students can normally expect a response within 24 hours, excluding weekends and University holidays. Students are responsible for checking their UARK email and the course Blackboard site regularly. Course announcements, including changes caused by a canceled class, will be sent through these channels.

Course Description

This course introduces machine learning from a mathematical and computational point of view. We study how learning algorithms represent data, fit models, evaluate predictions, and manage uncertainty. The course begins with Python, Git, and common data tools, then develops classical machine learning methods such as linear regression, logistic regression, clustering, nearest neighbors, support vector machines, decision trees, random forests, and boosting. The final third of the course connects these foundations to neural networks and modern deep learning, with implementation work oriented toward PyTorch, extending through convolutional and recurrent networks to variational autoencoders and generative models.

The course is designed for students who want to understand why machine learning methods work, how to implement them responsibly, and how to apply them to real data.

The course develops each method from two viewpoints at once: the **mathematics** of the model and the **engineering** of training it. Students should be able both to derive and justify a method and to run and diagnose its training process. This distinction carries into assessment: **exams test the mathematics** (derivations, objective functions, optimization, bias–variance, margins and kernels, backpropagation), while **quizzes test the training process** (data splits, feature scaling, learning-rate and epoch behavior, over/underfitting diagnosis, and metric choice). Both draw from the same weekly material.

Prerequisites

MATH 3083 Linear Algebra or MATH 3093 Abstract Linear Algebra, or approval of the department. Students should be comfortable with vectors, matrices, functions, and basic mathematical notation. Prior Python experience is helpful but not required.

Learning Outcomes

After completing the undergraduate version, students should be able to:

- Explain the difference between supervised, unsupervised, and deep learning methods.
- Use Python, Git, NumPy, pandas, scikit-learn, and introductory PyTorch tools for reproducible machine learning work.
- Implement and evaluate core algorithms for regression, classification, clustering, dimensionality reduction, and tree-based learning.
- Interpret model performance using train/test splits, cross-validation, loss functions, and classification/regression metrics.
- Complete a portfolio of homework assignments, with readable code, clear evaluation, and concise written explanation, spanning classical methods through neural networks and generative models.

Students enrolled in the graduate version will additionally be expected to:

- Derive or justify key algorithmic steps for selected models, including optimization or objective-function arguments.
- Compare algorithms at a higher level of abstraction, including assumptions, limitations, and failure modes.
- Meet higher standards for written explanation, reproducibility, and synthesis of mathematical and computational ideas.

Topics

- Python, Git, data preprocessing, and reproducible workflows; brief orientation to AI/LLM agents relative to classical ML.
- Linear regression, logistic regression, gradient descent, and regularization.
- Model evaluation, cross-validation, bias-variance tradeoff, and resampling.
- K-nearest neighbors, k-means clustering, PCA, t-SNE, and manifold-learning motivation.
- Support vector machines and kernel methods.
- Decision trees, random forests, gradient boosting, and ensemble methods.
- Neural networks, backpropagation, and PyTorch.
- Convolutional and recurrent neural networks.
- Variational autoencoders and generative models.

Textbook and References

There is no required textbook. The course materials are distributed in three places. The **course repository** on UARK GitLab holds the lecture notes, lab specifications, homework assignments, datasets, and this syllabus; you will fork it and push your own work there. Accessing it requires a UARK account and either an on-campus connection or the university VPN. **Slide decks** are posted on the course server, whose address is given in class, on Blackboard, and in the repository's `README`; if a slide and the notes ever disagree, the notes are authoritative. Recordings of the three asynchronous class meetings are posted on that same server. The lecture notes are also published openly on the course page. The course reading map in `readings.md` at the top of the repository

gives week-by-week references. The main references are Hastie, Tibshirani, and Friedman, *The Elements of Statistical Learning*; Bishop, *Pattern Recognition and Machine Learning*; Spiliopoulos, Sowers, and Sirignano, *Mathematical Foundations of Deep Learning Models and Algorithms*; and the official scikit-learn, pandas, NumPy, and PyTorch documentation. Some reference books may be available only through the library or instructor-provided excerpts, depending on copyright permissions.

Student Responsibilities

Students are responsible for preparing for class, participating in learning activities, monitoring course announcements, completing work by the stated deadlines, and asking questions when course expectations are unclear. Students should retain copies of submitted work and verify that files have been uploaded or pushed successfully. Technical difficulty should be reported to the instructor before the deadline whenever possible.

Course Work and Grading

The course grade is based on homework, weekly labs, Blackboard notes, quizzes, and three exams. There is no final project; all coursework beyond labs, quizzes, and exams, including the neural network and generative-model material in the final third of the course, is completed as homework.

Component	Weight	Description
Homework	35%	Eleven math and coding assignments connecting theory, derivations, and implementation, from linear regression through generative models. Each is worth 10% of the homework grade; the lowest score is dropped.
Weekly Labs	15%	Short coding and workflow assignments emphasizing practice, implementation, and reproducibility.
Blackboard Notes	10%	Twenty-three short discussion posts tied to lecture, coding days, and interpretation practice. Each is graded complete/incomplete; the three lowest are dropped, so the best twenty count. Graded for engagement and reasoning, not for a single correct answer. Substantive activity in the term-long Open Discussion forum can recover up to two further missed notes.
Quizzes	5%	Short weekly checks on the training/engineering process (splits, scaling, learning rate, over/underfitting, metrics).
Exams	35%	Three math-focused exams (Sep 16, Oct 14, Dec 2), roughly equal weight, covering derivations and the mathematics of the methods.

Letter grades use a stricter scale for the graduate section. The final grade will be no stricter than the following:

Letter	4990V (Undergraduate)	5990V (Graduate)
A	80–100	90–100
B	70–79	80–89
C	60–69	70–79
D	50–59	60–69
F	Below 50	Below 60

Labs and Homework

Labs and homework are submitted by pushing to your own fork of the course repository on UARK GitLab; the repository's `README.md` walks through forking, cloning, and the per-assignment workflow. Blackboard notes are submitted in Blackboard. Assignments must be pushed or uploaded before the posted deadline.

Homework is due Monday; labs are due Wednesday (see `course_schedule.pdf` for exact dates and holiday adjustments). Homework submitted within 24 hours after the deadline receives a 10 percentage-point deduction; homework submitted 24–48 hours late receives a 20 percentage-point deduction. Homework more than 48 hours late receives no credit unless an exception has been approved for an excused absence or emergency. The lowest homework score is dropped as described above.

Labs have a firm deadline and are not accepted late except when an excused absence or emergency prevents timely completion. Students seeking an exception should contact the instructor in writing before the deadline when possible, or as soon as reasonably possible afterward. An approved extension establishes a new deadline that applies to that student.

Blackboard Notes and Open Discussion

The Blackboard Notes category (10%) has two parts: **22 graded prompts** tied to individual class meetings, and one ungraded **Open Discussion** forum that runs all term.

Blackboard notes (the graded prompts)

A note is a short written response to a prompt tied to that day's class. It asks you to interpret evidence, diagnose a result, or explain a choice in your own words. Notes are *not* derivations and *not* code — if the work already lives in a lab or homework, do not re-upload it here.

- **Where.** Each prompt has its own discussion forum in Blackboard.
- **Post first, then read.** The forums are set so that you must post your own response before you can see anyone else's. The point is to capture your own thinking while the class is fresh. After you post, you are encouraged to read how classmates approached the same question — several prompts have more than one defensible answer. Replying is welcome but not required, and only your own post is graded.
- **How many.** Twenty-three are assigned; the **three lowest are dropped**, so your best twenty count.
- **How graded. Complete or incomplete.** A note is complete when it is on time, answers every part of the prompt, and refers to at least one concrete detail — a number, a feature of a plot, a parameter value, or a specific result from your own code. A response that is on topic but entirely generic does not receive credit.

- **When due.** Lecture prompts are due before the next class meeting; prompts attached to an asynchronous session are due the following Sunday night. Exact dates are posted with each prompt.
- **Late work.** Late notes are not ordinarily accepted. The three drops exist to cover an illness, a missed class, or a hard week without a separate request.

Open Discussion (ungraded, always open)

Open Discussion is a single forum that stays open for the whole semester, and **everything in the course is fair game there** — lectures, labs, homework, exam preparation, setup and tooling problems, a result you found interesting, or continuing an argument about a graded prompt after its forum has closed. If you are stuck on an assignment, ask. Working a problem out together is how most people learn this material. Posting rules are pinned at the top of the forum.

Open Discussion carries no points of its own, but it can **recover missed notes**. If you miss more than the three dropped notes, up to **two additional** missing notes may be replaced by substantive contributions here — so at most five of the twenty-three can be covered this way. A substantive contribution is a post or reply that materially helps someone: a specific answer, a worked explanation, a resource with a sentence on why it is useful. "Thanks," "same question," and one-line agreement do not count. If you want contributions counted, email me before the last week of class with links to the posts you are claiming.

The course collaboration policy applies here exactly as written elsewhere in this syllabus: discuss freely, but the work you submit must reflect *your own* understanding, and you should acknowledge classmates whose help was substantial. The only thing worth avoiding is short-circuiting that — posting a finished solution or a block of code that completes an assigned problem takes the thinking away from everyone who reads the thread afterward, and a submission copied from it is not your own work. Explain the idea, point at the relevant section of the notes, or help debug the error instead.

Attendance and Participation

Education at the university level requires active involvement in the learning process. Students are expected to attend class and participate in lectures, labs, quizzes, discussions, and other learning activities. Attendance itself is not a separate graded category, but work completed in class may contribute to the lab, quiz, or Blackboard Notes categories.

Illness, a family crisis, a required University-sponsored activity, religious observance, jury duty or a court subpoena, military duty, and comparable circumstances may make attendance impossible. Students should notify the instructor in writing before the absence when possible and make timely arrangements for missed work. The instructor may request appropriate documentation. Students remain responsible for obtaining announcements and course material covered during an absence. See the University's attendance policy for additional information.

Make-Up Work

An exam or quiz missed because of an excused absence may be replaced by an equivalent make-up assessment or, for a quiz, omitted from the category calculation at the instructor's discretion. Students must contact the instructor before the assessment when possible, or within 48 hours after the absence when advance notice is not possible. Unexcused missed exams and quizzes receive a

zero. Work affected by a University closure or instructor cancellation may be submitted or made up without penalty under the cancellation policy below.

Exams and Quizzes

There are three exams, held on **Wednesday, September 16**; **Wednesday, October 14**; and **Wednesday, December 2**. Exams are cumulative in spirit and focus on the mathematics of the methods. Short quizzes are given in most weeks and focus on the training and engineering process introduced in that week's lab. Blackboard notes are short discussion or reflection assignments tied to lecture, coding days, and interpretation practice. Undergraduate and graduate students take the same exams and quizzes; graduate students are held to the stricter grade scale above and to additional expectations described under Learning Outcomes.

Unless an exam announcement states otherwise, students may use one student-prepared, handwritten, 8.5-by-11-inch reference sheet. Calculators, phones, smart watches, laptops, tablets, and other electronic devices may not be used during exams. Quizzes are open notes unless announced otherwise. Students must show the main steps supporting answers on written assessments.

Use of AI Tools

Students may use AI tools as learning aids, debugging aids, and writing-support tools when permitted by the assignment instructions. Students remain responsible for understanding and explaining submitted work. Any substantial AI assistance used to produce code, analysis, or writing should be acknowledged according to the assignment instructions.

Academic Integrity

The University of Arkansas provides opportunities for study and research in an environment that depends on intellectual honesty and individual integrity. Students must be familiar with and follow the University's Academic Integrity Policy. Collaboration is encouraged when an assignment permits it, but submitted work must reflect the student's own understanding and must acknowledge external sources, code, collaborators, and substantial assistance. Students who are uncertain whether a form of collaboration or tool use is permitted should ask before submitting the work.

Classroom Conduct and Electronic Devices

Students and the instructor share responsibility for maintaining an environment that supports learning. Everyone is expected to communicate respectfully and keep disruptive behavior to a minimum. Side conversations, repeated late arrival or early departure, and use of devices for activities unrelated to the course interfere with that environment. Phones should be silenced and put away during class. Laptops and tablets may be used for course activities, particularly coding labs, but may be restricted during particular discussions or assessments. Disruptive conduct may result in a request to leave class and may be referred to the Office of Academic Integrity and Student Conduct.

Intellectual Property and Course Materials

Instructor-created notes, slides, videos, exams, quizzes, solutions, and other course materials are provided for students' personal educational use in this course. Students may not sell, publicly post, or distribute these materials without written permission. This restriction does not prevent students from sharing their own notes or working with classmates in ways permitted by the course's academic integrity and collaboration rules.

Class Cancellation, Inclement Weather, and Emergencies

Class will meet when the University is open. Students traveling from off campus should make decisions that protect their personal safety and should contact the instructor when weather prevents safe travel. If the University delays, dismisses, or closes because of weather, class may be canceled or conducted remotely at the scheduled time. Synchronous attendance will not be required while campus is closed; any synchronous remote class will be recorded and made available through Blackboard. Work due during a University closure may be made up without penalty.

If the instructor must cancel class because of illness or another emergency, an announcement will be sent by UARK email and posted to Blackboard. Missed instruction may be replaced with a substitute instructor, remote class, recorded lecture, reading, assignment, or other instructional material. Instructions will be available before the scheduled class time when feasible. Students will have an opportunity to make up work directly affected by the cancellation. Emergency guidance is available at `https://tips.uark.edu/emergency-preparedness/`.

Accessibility

Under University policy and federal and state law, students with documented disabilities are entitled to reasonable accommodations that provide equal access to the course. Students seeking accommodations should contact the Center for Educational Access (CEA). After CEA approves an accommodation, the student should meet with the instructor as early as possible, preferably during the first week of class or within one week of registration with CEA, to discuss its implementation. Accommodations are not applied retroactively.

Student Support and Well-Being

Students can find academic support through the Student Success Center and mathematics support through the Mathematics Resource and Teaching Center. Health and wellness resources are listed in the University catalog. Students facing circumstances that affect their participation are encouraged to contact the instructor early so that available course and University support can be identified.

Grade Questions and Disputes

Scores will be posted in a timely manner. Students should review posted scores and bring a possible grading or recording error to the instructor in writing within one week of posting. The request should identify the specific item and explain the concern. All scores posted before Dead Day will be treated as accurate unless a possible error is reported before Exam 3. This review period does not limit rights provided by the University's formal grade-appeal procedures.

Tentative Schedule

The week-by-week schedule is maintained in `syllabus/course_schedule.pdf` in the course repository, and is also posted on the course page. The schedule is tentative and may be adjusted to support student learning and official university calendar changes.

Syllabus Changes

This syllabus and the course schedule may be revised when necessary. Material changes to policies, deadlines, or course requirements will be announced through UARK email and Blackboard. Changes will be applied consistently and will not be made retroactively to a student's disadvantage.