COMPSCI 589 – Machine Learning
Syllabus

Spring 2024

Instructor:	Bruno Castro da Silva	**Time:**	Tue & Thur (10:00 – 11:15)
Email:	bsilva@cs.umass.edu	**Place:**	Thompson Hall, Room 102

1 About the course

This course will introduce core machine learning models and algorithms for classification, regression, clustering, and dimensionality reduction. On the theory side, the course will cover the mathematical foundations underlying the most commonly-used machine learning algorithms. It will focus on understanding models and the relationships between them. On the applied side, the course will focus on effectively using machine learning methods to solve real-world problems with an emphasis on model selection, regularization, design of experiments, and presentation and interpretation of results. The course will have assignments that involve both mathematical problems and implementation tasks.

Broad topics covered in this course will include classification algorithms in general, decision trees, random forests, probabilistic models, Naive Bayes methods, various ensemble meta-algorithms (such as bagging and boosting), gradient-based techniques, linear regression, logistic regression, neural networks, convolutional neural networks and deep learning, unsupervised learning and clustering algorithms, k-means, hierarchical clustering, and dimensionality reduction techniques.

In this course, each voice in the classroom has something of value to contribute. Please take care to respect the different experiences, beliefs, and values expressed by students and staff involved in this course. My colleagues and I support UMass' commitment to diversity, and welcome individuals regardless of age, background, citizenship, disability, sex, gender, gender identity, sexual orientation, education, ethnicity, family status, geographical origin, language, military experience, political views, race, religion, socioeconomic status, and work experience.

2 Website

This course's syllabus will be hosted here. Homework assignments and other material will be posted on Canvas. Lectures will be recorded; recordings (along with .pdf slides) will also be on Canvas.

3 Class

Classes will be held on Tuesdays and Thursdays from 10:00 to 11:15 am in Thompson Hall, Room 102.

4 Textbooks

The course has no mandatory textbook. However, you may find the following (freely available) books helpful:

- Machine Learning: a Probabilistic Perspective, by Kevin Patrick Murphy.
- The Elements of Statistical Learning, by Hastie, Tibshirani and Friedman.
- An Introduction to Statistical Learning, by James, Witten, Hastie and Tibshirani.

5 Required background

While this course has an applied focus, it still requires appropriate mathematical background in probability and statistics, multivariate calculus, linear algebra, and programming. The official prerequisites for undergraduate students can be found here. Graduate students can check the descriptions for these courses to verify that they have sufficient mathematical background for COMPSCI 589. The following references can provide a useful review:

- Probability Theory
- Linear Algebra and Matrix Calculus
- Matrix Cookbook
- Optimization: Any calculus textbook.

6 Attendance policy

Please make every effort to attend and actively participate in each class meeting. However, you are free to choose not to attend "regular" class meetings (i.e., classes where no assignments or exams are given) if the circumstances warrant. We do understand that all students are capable of autonomously deciding how to best allocate their time. *However, notice that you are responsible for classes you miss.* In the event that you miss class, it is your responsibility to obtain the missed notes from a classmate and for picking up any handouts distributed in class by the instructor. The instructor will not provide class notes or condense the entire contents of a missed class period in an e-mail message or a five-minute summary. *Important: please see Section 11 of this document for more details on this course's late policy regarding late homework submissions and missed exams.*

7 Teaching Assistants

The teaching assistants (TAs) this semester will be:

- Khoshrav Doctor (kdoctor@cs.umass.edu)
- Aline Weber (alineweber@umass.edu)
- Andrew Yuan (awyuan@umass.edu)
- Prithviraj Tarale (ptarale@umass.edu)
- Erfan Entezami (eentezami@umass.edu)

8 Office Hours

Office hours will be held according to the following schedule (starting on 02/12), except *(1)* on holidays; *(2)* if the UMass official schedule follows a different day of the week; or *(3)* when noted otherwise.

	Monday	**Tuesday**	**Wednesday**	**Thursday**	**Friday**
8:50am–9:50am		Prithviraj			
9:00am–10:00am	Erfan		Erfan		Prithviraj
11:15am–12:30pm		Bruno		Bruno	
12:00pm–1:00pm		Khoshrav (*)			Khoshrav
1:00pm–2:00pm	Aline		Aline		
2:30pm–3:30pm		Andrew		Andrew	

If one of the TA's or the instructor needs to reschedule their office hours, an official announcement will be made on Canvas. **All office hours will be in person unless noted otherwise.** The Zoom link below will be used in cases when there is a particular reason a student cannot attend in person—for example, if they are sick or traveling—or if they made arrangements with the TA, in advance, to schedule a particular time to be helped via Zoom.

Office Hours – Locations:

- TA's in-person office hours: LGRT T220 (except for those marked with an (*) in the table above. These will take place in LGRT T222).
- Remote office hours: on Zoom.
- **Until noted otherwise**, Prof. da Silva's office hours will be held immediately after each lecture. He will remain in the classroom for 1 hour and 15 minutes, post-lecture, helping students as needed.

9 Grading

Your grade will have three components:

1. **Homework Assignments** (55%): There will be frequent homework assignments, both written and programming. All assignments will have an equal weight.
2. **Midterm exam** (30%).
3. **Project** (15%): Given the wide range of possible real-life applications of machine learning, it is critical that we study how to implement, fine-tune, and deploy these algorithms in practice. Further details will be available when the project is assigned, after the most relevant course material has been covered.

A cumulative grade in $[90\% - 100\%]$ will be an A- or A, $[75\%, 90\%)$ will be a B-, B, or B+, $[65\%, 75\%)$ will be a C-, C, or C+, and $[55\% - 65\%)$ will be a D or D+. Course grades will be curved only in students' favor (that is, these thresholds may be lowered, but a grade of 90% will not be lower than an A-). **Some extra credit opportunities may be given. Your grade may be reduced by any amount at the instructor's discretion due to inappropriate behavior, such as academic dishonesty.**

10 Re-grading policy

Errors in grading of assignments and exams can occur despite the best efforts of the course staff. If you believe you've found a grading error, complete an online re-grade request form via Gradescope. Re-grade requests must be submitted no later than one week after the assignment is returned. Note that re-grading may result in your original grade increasing or decreasing as appropriate.

11 Late Policy

- Deadlines in this course are **strict**. A submission one minute after the deadline will receive zero credit. You are strongly encouraged to submit hours before any deadline.
- Having said that, to allow some flexibility to complete assignments (*homeworks only*) given other constraints, you have a total of *seven* free late days that you can choose to use when submitting a homework. You will be charged one late day for handing in an assignment within 24 hours after it is due, two late days for handing in an assignment within 48 hours after it is due, etc. Your assignment is considered late if either the written or code portions are submitted late. The late homework clock stops when both the written and code portions are submitted. After you have used up your late days, late homework will not count for credit except in special circumstances (e.g., illness documented by a doctor's note).
- All exams must be taken at the scheduled time unless *(1)* there is a documented conflict and arrangements have been made with the instructor before the exam; or *(2)* you have a medical emergency and you bring proof of such to the instructor before final grades for the given exam are computed. In any other case (unless those covered by the University's Academic Regulations), missing an exam will result in a grade of "F" for that exam.

12 Pass/Fail & SAT/Fail

- If you are an undergraduate student, Pass/Fail is requested through the university.
- If you are a graduate student, at some time near the end of the semester (likely around the last day of class), you will be given the option to take the class SAT/Fail rather than for a letter grade. If you plan to take the course SAT/Fail, keep an eye out for an email (or a message on Canvas) from me around the end of the semester with instructions for requesting SAT/Fail. If you elect SAT/Fail, you will earn a SAT grade if your letter grade would have been a C or higher, and you will receive an F if your letter grade would have been lower.
- The above conditions *do not* hold for students with an academic honesty violation. In these cases, the requests described in this section are disallowed and/or un-approved.

13 Disability Services

The University of Massachusetts is committed to providing an equal educational opportunity for all students. If you have a documented physical, psychological, or learning disability on file with Disability Services, you may be eligible for academic accommodations to help you succeed in this course. If you would like to register with Disability Services, please visit their website or their office (161 Whitmore Administration Building; phone (413) 545–0892). Finally, if you have a documented disability that requires an accommodation, please notify me within the first two weeks of the semester so that we can make appropriate arrangements.

14 Cheating

- Cheating will not be tolerated. Assignments may include instructions about what forms of collaboration are allowed, if/when relevant.
- Copying answers or code from external sources (books, web pages, etc.), from other students, or from solutions to assignments from previous years is *always* considered cheating. Note that, according to the new UMass Academic Honesty Policy, the use of AI text generators (such as **ChatGPT**) is **prohibited**. To emphasize: no detectable copying is acceptable, even, e.g., copying a single sentence from an outside source. Sharing your code or solutions with other students is also considered cheating.
- The College of Information and Computer Sciences explicitly forbids any redistribution (including publicly available posting on an internet site) of any CICS course materials (including student solutions to course assignments, projects, exams, etc.) without the express written consent of the instructor of the course from which the materials come. Violations of this policy will be deemed instances of "facilitating dishonesty" (since a student making use of such materials would be guilty of plagiarism) and therefore may result in charges under the Academic Honesty Policy.
- Since students are expected to be familiar with this policy and the commonly accepted standards of academic integrity, ignorance of such standards is not normally sufficient evidence of lack of intent.
- **All instances of cheating will be reported to the university's Academic Honesty Board.** Any detected cheating will result either *(i)* in a grade of -100% on the assignment for all students involved (negative credit); or *(ii)* a grade of F in the course. The instructor will decide at their discretion which of these possible resolutions is more appropriate.

15 COVID-19 and Face Covering Policy

Students and course staff in COMPSCI 589 are expected to do their part to slow the spread of COVID-19 and minimize the risk of illness for all community members. It is important to remember that community members may be vulnerable or live with vulnerable individuals.

Masks are welcome on campus, and we encourage everyone to respect the choices that individuals make about their own masking. Masking is encouraged during the **first few weeks** of the Spring semester.

For general information, please refer to the official UMass FAQ regarding COVID-19 and masking.

Students and course staff in COMPSCI 589 must comply with all university policies regarding COVID-19.

- **If you test positive for COVID-19**, *you need to self-isolate*: don't come to class! Individuals who test positive for COVID-19 are required to isolate for a minimum of *five days* before returning to class. They should then continue to wear a mask for an additional *five days*.
- **If you were exposed to COVID-19**, you need to wear a high-grade mask (such as KN95, KF94, or N95 for *ten days*). You also need to get tested at least *five full days* after your exposure.

If you are in any doubt, please do not come to class—you can request an excused absence from any course meeting by sending the instructor an email; you will not be penalized for any missed quizzes, discussion