Syllabus

Course Info

- **Instructors**: Nina Balcan, Leila Wehbe
- **Education Associate:** Daniel Bird dpbird@andrew.cmu.edu
  Joshmin Ray joshminr@andrew.cmu.edu
- **Meetings**:
  - **10-315:** MWF, 12:20 PM - 01:40 PM
  - For all sections, lectures are on Mondays and Wednesdays.
  - Occasional recitations are on Fridays and will be announced ahead of time.
- **Piazza:** https://piazza.com/class/kjvtxt3vqtebc
- **Gradescope:** https://www.gradescope.com/courses/228169
- **Video:** https://canvas.cmu.edu/courses/21501/external_tools/269
- **Schedule:** https://docs.google.com/spreadsheets/d/19MU6chQjTB-a1VLVbcf7Xiv1iuHKFFWT5GIj3WlCyFg/edit?usp=sharing
- **Office Hours Queue:** https://cmu.ohqueue.com
- **Course Calendar:** https://calendar.google.com/calendar/u/1?cid=Y18yZ3FzcW1iNWNoYXBkNml2bDZoYjNtZHE0c0Bncm91cC5jYWxlbmRhci5nb29nbGUuY29t

1. Course Description

Machine learning is a subfield of computer science with the goal of exploring, studying, and developing learning systems, methods, and algorithms that can improve their performance with learning from data. This course is designed to give undergraduate students a one-semester-long introduction to the main principles, algorithms, and applications of machine learning and is specifically designed for the SCS undergrad majors. After completing the course, students will be able to:

- Select and apply an appropriate supervised learning algorithm for classification problems and understand its underlying assumptions (e.g., naive Bayes, perceptron, support vector machine, logistic regression).
- Select and apply an appropriate supervised learning algorithm for regression problems and understand its underlying assumptions (e.g., linear regression, ridge regression).

- Recognize different types of unsupervised learning problems, and select and apply appropriate algorithms (e.g., clustering, linear and nonlinear dimensionality reduction).
- Work with probabilities (Bayes rule, conditioning, expectations, independence), linear algebra (vector and matrix operations, eigenvectors), and calculus (gradients) to derive machine learning methods such as linear regression, naive Bayes, and principal components analysis.
- Understand machine learning principles such as model selection, overfitting, and underfitting, and techniques such as cross-validation and regularization.
- Implement machine learning algorithms such as logistic regression via stochastic gradient descent, linear regression (using a linear algebra toolbox), perceptron, or k-means clustering.
- Run appropriate supervised and unsupervised learning algorithms on real and synthetic data sets and interpret the results.

Key Topics

This course covers the core concepts, theory, algorithms and applications of machine learning. We cover supervised learning topics such as classification (Naive Bayes, Logistic regression, Support Vector Machines, neural networks, decision trees, boosting) and regression (linear, nonlinear) as well as unsupervised learning (MLE, MAP, clustering, PCA, dimensionality reduction).

2. Prerequisites

The formal prerequisites are at least a "C" in each of these courses:

- 15-122 - Principles of Imperative Computation
- 21-127 - Concepts of Mathematics (or 21128, 15151)
- 21-325 - Probability (15359, 36217, 36218, 36219, 36225)
- 21-241 - Matrices and Linear Transformations

3. Recommended Textbooks

- Machine Learning, Tom Mitchell.(optional)
- Pattern Recognition and Machine Learning, Christopher Bishop. available online , (optional)
- Machine Learning: A Probabilistic Perspective, Kevin P. Murphy, available online, (optional)

The core content of this course does not exactly follow any one textbook.

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:

- 50% Homework Assignments
- 15% Midterm Exam 1
- 15% Midterm Exam 2
- 20% Final Exam

Midterms and Final Exam

You are required to attend the midterms and final exams. The midterm exams will be given in class. The final exam will be scheduled by the registrar sometime during the official final exams period. Please plan your travel accordingly as we will not be able accommodate individual travel needs (e.g. by offering the exam early).

If you have an unavoidable conflict with an exam (e.g. an exam in another course), we will be releasing a conflict form a few weeks before the examination, please fill out these forms at your earliest convenience so that we can accommodate you.

No electronic devices are allowed during the exam. Unless otherwise noted, all exams are closed-book.

Homework

The homeworks may be divided into two components: programming and written. The programming assignments will ask you to implement ML algorithms; they emphasize understanding of real-world applications of ML, building end-to-end systems, and experimental design. The written assignments will focus on core concepts, “on-paper” implementations of classic learning algorithms, derivations, and understanding of theory.

Recitations

Attendance at recitations (Friday sessions) is not required, but strongly encouraged. These sessions will be interactive and focus on problem solving. If you are unable to attend one or you missed an important detail, feel free to stop by office hours to ask the TAs about the content that was covered. Of course, we also encourage you to exchange notes with your peers.

5. Technologies

We use a variety of technologies:

Piazza

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

Gradescope

We use Gradescope to collect PDF submissions of **open-ended questions** on the homework (e.g. mathematical derivations, plots, short answers). The course staff will manually grade your submission, and you'll receive personalized feedback explaining your final marks.

You will also submit your code for **programming questions** on the homework to Gradescope. After uploading your code, our grading scripts will autograde your assignment by running your program on a VM. This provides you with immediate feedback on the performance of your submission.

Regrade Requests: If you believe an error was made during manual grading, you'll be able to submit a regrade request on Gradescope. For each homework, 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 homework policy

You receive 4 total grace days **for use on any homework assignment**. We will automatically keep a tally of these grace days for you; they will be applied greedily. No assignment will be accepted more than 2 days after the deadline. Submitting late with no grace days remaining will lead to the following penalties:

- 50% after 1 day
- 25% after 2 days
- 0 after that

This has two important implications: (1) you may not use more than 2 grace days on any single assignment (2) **you may not combine grace days with the late policy above to submit more than 2 days late.**

For example, you cannot use two grace days for an assignment and then submit on the third day for 50% credit. You would instead receive a score of zero for that homework. If you instead only have one grace day remaining and submit an assignment on the second day, you will use that grace day and take a 50% penalty.

All homework submissions are electronic (see Technologies section below). As such, lateness will be determined by the latest timestamp of any part of your submission. For example, suppose the homework requires submissions to both Gradescope and Autolab – if you submit to Gradescope on time but to Autolab 1 minute late, your entire homework will be penalized for the full 24-hour period.

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 permitted this semester. In order to successfully audit the course, at least 50% of the homeworks must be completed, but exams will not be required.

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 Joshmin as early in the semester as possible. Joshmin 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. Additionally please forward details of accommodations to joshminr@andrew.cmu.edu.

7. Academic Integrity Policies

Read this carefully!

(Adapted from Roni Rosenfeld's 10-601 Spring 2016 Course Policies.)

Collaboration among Students

- The purpose of student collaboration is to facilitate learning, not to circumvent it. Studying the material in groups is strongly encouraged. It is also allowed to seek help from other students in understanding the material needed to solve a particular homework problem, provided no written notes (including code) are shared, or are taken at that time, and provided learning is facilitated, not circumvented. The actual solution must be done by each student alone.
- The presence or absence of any form of help or collaboration, whether given or received, must be explicitly stated and disclosed in full by all involved. Specifically, each assignment solution must include answering the following questions:
    1. Did you receive any help whatsoever from anyone in solving this assignment? Yes / No.
        - If you answered 'yes', give full details: _____
        - (e.g. "Jane Doe explained to me what is asked in Question 3.4")
    2. Did you give any help whatsoever to anyone in solving this assignment? Yes / No.
        - If you answered 'yes', give full details: _____
        - (e.g. "I pointed Joe Smith to section 2.3 since he didn't know how to proceed with Question 2")

3. Did you find or come across code that implements any part of this assignment ? Yes / No. (See below policy on “found code”)
    - If you answered ‘yes’, give full details: _____
    - (book & page, URL & location within the page, etc.).

- If you gave help after turning in your own assignment and/or after answering the questions above, you must update your answers before the assignment’s deadline, if necessary by emailing the course staff.
- Collaboration without full disclosure 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. You must solve the homework assignments completely on your own. We will be actively monitoring your compliance. Collaboration with other students who are currently taking the class is allowed, but only under the conditions stated above.

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 proactively 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 solutions 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.