Spring 2022
24-787 Machine Learning and Artificial Intelligence for Engineers

Instructor	Levent Burak Kara (lkara@cmu.edu)	
Class	MW 12:20 PM – 2:10 PM TEP 1308, Zoom for recordings (see Canvas for links)	
Recitation	F 3:35 PM – 4:25 PM TEP 1308, Zoom for recordings (see Canvas for links)	
Course Assistants	[TA] Akshay Hinduja <ahinduja@andrew.cmu.edu>	Office Hr: F – 2:30PM
	[TA] Jiacheng Zhu <jzhu4@andrew.cmu.edu>	Office Hr: F – 12:30PM
	[CA] Wangchuan Feng <wangchuf@andrew.cmu.edu>	Office Hr: W – 9:30AM
	[CA] Viraj Ranade <notifications@instructure.com>	Office Hr: TR – 11:00AM
	[CA] Dijing Zhang <dijingz@andrew.cmu.edu>	Office Hr: M – 10:00AM
	[CA] Shivani Saboo <ssaboo@andrew.cmu.edu>	Office Hr: T – 10:00AM
		Office hours will be 50mins and on Zoom (see Canvas for links)
Recommended Text Books	Introduction to Statistical Machine Learning. Masashi Sugiyama, 2016. Pattern Recognition and Machine Learning. Christopher M. Bishop, 2006. Reinforcement Learning: An Introduction. R. S. Sutton, A. Barto, 2020. Deep Learning by Ian Goodfellow, 2016: Available online (https://www.deeplearningbook.org) There is no single textbook that covers all of the topics we will discuss in this course. Hence, there will be no required textbook and instead you will be provided with copies of lecture notes. The recommended texts referenced above may be useful for a deeper understanding of the course subjects.	
Website	https://canvas.cmu.edu https://piazza.com	

Description:

This course provides an introduction to the fundamental methods and algorithms at the core of modern machine learning. It also covers theoretical foundations as well as essential algorithms and practical techniques for supervised and unsupervised learning.

Topics covered (tentative):

- Introduction to Machine Learning and Supervised Learning
- Regression
- Parametric/Non Parametric Learning
- Discriminative and Generative Algorithms
- Naive Bayes, Non-linear Classifiers
- Feature Engineering/Representation
- Ensemble Methods
- Support Vector Machine (SVM)
- Unsupervised Learning and Clustering Algorithms
- Principal Component Analysis, Independent Component Analysis
- Neural Networks

- Training, Testing and Evaluation
- Reinforcement Learning

Prerequisites

- Familiarity with the basic linear algebra.
- Familiarity with the basic probability theory.
- Basic computer programming skills and scientific computing.

Learning Outcomes

- Understand the fundamental problems addressed by machine learning and artificial intelligence.
- Gain hands-on experience in designing and implementing various machine learning algorithms.
- Understand the mathematical foundations for machine learning algorithms including linear algebra, probability, statistics, and optimization.
- Become experienced in formulating data-driven approaches to problems and communicating these solutions through implemented algorithms and write-ups.

Recitations

Since many of you may not have familiarity with Python and its libraries, there will be weekly recitations to cover your needs for python knowledge, programming and its integration with machine learning. Some of recitations are designed to teach you how to implement algorithms and theoretical machine learning into programs. We highly recommend you participate in the recitations.

Python Programming

We will use the Python programming language for all assignments in this course. Python is a great general-purpose programming language on its own, but with the help of a few popular libraries (numpy, scipy, matplotlib) it becomes a powerful environment for scientific computing.

Assignments

Most of the assignments in this course involve writing computer programs. In a typical assignment, you will implement a machine learning technique from the lecture and use it to solve a sample problem. You will be graded on how well your computer program works. Therefore you should carefully implement, test and debug each program. Remember, just because the compiler gives no error messages, does not mean the program works as it should. In addition to submitting your code, you will typically be required to annotate and comment your program. To make programming, reporting, commenting, plotting and submitting your program easier, we will use **iPython notebooks (Jupyter Notebooks).** All assignments will be in **Python**.

Getting Help from Teaching Assistants

- A TA/Grader is not supposed to debug the code for a student.

- A TA/Grader cannot go beyond the allocated time during office hours.

- A TA/Grader will only answer specific questions about the problem and not generic questions like "How do I solve this problem". The office hours are usually not helpful unless you come prepared with specific questions.

- A TA/Grader may or may not respond to emails in the weekend. Do not rely on TA availability on the weekend.

- A TA/Grader can meet you outside of the office hour depending on if he/she has some time.

Submission Procedures:

- Submission procedures are explained for each assignment on its first page. If you want to take photos of your assignment, just make sure that combine all the jpg into a **single pdf file** and make it clear. Unorganized and illegible submissions will be penalized (see below). Take care in arranging your illustrations, written solutions, photos.
- Only submit the required files that you edit. All other files, including the ones we provide as supporting functionality, are unnecessary because we already have copies of them. Unless we tell you otherwise, make sure that you only submit the file you edited/changed.
- Do not scan your hand-written solutions in the highest resolution. Having us download 200MB worth of scanned files from Canvas is unreasonable. Exercise good judgment. You can set the resolution when scanning. Make sure your submission zip file is less than 20MB. That's enough for all the assignments.
- Name your files logically and clearly. Add a readme.txt with your submissions. It's hard to open each ill-named file one by one and find your solution.
- A failure to comply with these procedures will result in grade penalties that will be proportional to the severity of your non-compliance.

Grading:

Assignments: 100% Total = 100%

Grading Scale:
90%-100% = A
80%-90% = B
70%-80% = C
60%-70% = D
0%-60% = R

The grading of the assignment will be out of 100 points, but the final score will be multiplied by 0.95. Thus, 95% of the points you attain is contingent on the correctness of your answers. 5% is reserved for organizational skills, quality of work, easy accessibility and legibility. The extra credits as usual would be added to the final score.

Late Submission Policy:

You have the option to submit your assignment after the announced deadline (started late, had to travel, computer malfunction etc.). The following policy will apply:

- Submission after the deadline, but within the first 2 hours of the deadline: Your nominal assignment grade is multiplied by 0.95.

- Submission after the deadline, but within the first 24 hours of the deadline: Your nominal assignment grade is multiplied by 0.80.

- For further delays, 10% will be deducted for additional 24-hour blocks. For example, if you submit your assignment 30 hours passed the deadline, the multiplier is 0.70. 50 hours passed the deadline; the multiplier is 0.60 etc.

A note on collaboration:

For your assignments, we encourage you to collaborate with each other. However, some of you may have differing ideas of what appropriate collaboration means.

Here are some guidelines for you when it comes to collaboration:

1. Any text or graphics contained in your report should be solely your work. You should not use anything that was created by someone else.

2. Any code that you submit should be entirely written by you (excluding, of course, the framework we provide you).

3. To reiterate (1) and (2), you should not be cut-and-pasting from someone else, sending work via email to someone else, or having someone else type the work for you.

Here are some forms of good collaboration:

1. Discussing the implications of results with your classmates. "Hey, my neural network has worse results than my decision trees, but that doesn't make intuitive sense to me. What do you think?" "Yeah, my NN is better than my DT. Perhaps your code is doing <insert statement about the code> incorrectly?"

2. Asking for help conceptualizing something. "I don't get the purpose of the Kernel as it relates to SVM. Do you know what it means?" "Yes, it means..."

3. Discussing theorems, principles, equations, etc. "Hey, in the proof on Problem 2a, I tried using the definition of the dot product; did you do something similar?" "Yes, I did, and I noticed that the fact that cosine is an even function make the two sides of the equation equivalent."

4. Debugging code. "My code is giving me this error. Can you come take a look?" "Sure...oh yeah, I got that error as well. You can get rid of it by using this function instead."

If you are unsure of whether or not you are participating in inappropriate collaboration, talk to the TA. We reserve the right to deduct points if cheating is suspected.

Here is a useful link to CMU's policy on the matter:
https://www.cmu.edu/policies/student-and-student-life/academic-integrity.html

A note on Web sites providing solutions to homework problems such as Assignment Geek, Bartleby, Chegg, etc.]:

First, please note that the use of such solution services during exams or assignments is not allowed and is grounds for course failure and additional penalties (see https://www.cmu.edu/policies/student-and-student-life/academic-integrity.html). Use of these services is also forbidden when preparing homework solutions; posting homework questions to such services or accessing solutions to these problems will result in zero credit for the entire homework assignment.

Second, there is a broader concern to be aware of. Students have greatly increased access to homework solutions and online services are even available to provide solutions. Great caution should be used in accessing and using these materials in preparing homework. Homework is an opportunity to challenge yourself to apply the course material in a low-stakes assessment. An over reliance on these resources to complete an assignment will likely impede your learning of the material. The most significant challenge students face in these courses is typically the selection of a method for a physical scenario (e.g., what assumptions to use). If a solution is reviewed prior to sufficiently attempting a problem independently, that aspect of problem solving is not practiced. If you choose to use these resources in completing your homework, it is recommended that you review them after completing your attempt. If you have challenges in completing the homework, you are strongly advised to seek help during office hours in

order to assist you in identifying and addressing any possible knowledge gaps or for us to provide additional tips on the solution process.

Cheating:

In the event of a student cheating on any assignment / exam in this course, unless extraordinary circumstances prevail, I plan to impose a penalty of failure for the course and, as required, will report it to the university as a violation of academic integrity.

Course Logistics

- We will be using Canvas for the announcements, assignments, lecture notes, recitation files, and grading.
- Students should use Piazza for posting queries and finding team members for projects.
- You are encouraged to resolve any issues and queries during the Office hours.
- You are encouraged to check the Announcement section in Canvas periodically.
- Office hours will be spread across the week to better accommodate different schedules of all enrolled students.
- Polls will be taken on a regular basis to collect class feedback on lectures and recitations.

Take care of yourself:

This year is unlike any other. We are all under a lot of stress and uncertainty at this time. Attending Zoom classes all day can take its toll on our mental health. 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. 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 almost always 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 me know. Thank you, and have a great semester.