90-803 Machine Learning Foundations with Python

Spring 2024 (A2 and B2)

Instructor

Prof Gongora-Svartzman (Prof. GS)
Email: ggongora@cmu.edu
Office Hours: Look on CANVAS for more updates

Teaching Assistants

Look on CANVAS for more updates.

Course Description

Machine Learning (ML) is a subfield within artificial intelligence aimed at having algorithms that will mimic "intelligent human behavior". As AI has gained popularity over the past decade - and so has ML - and more healthcare and public policy applications nowadays have found it beneficial to incorporate ML into their decision-making process. This course is designed to give students a solid grounding in modeling and tools employed in machine learning and applied to healthcare and public policy scenarios. The emphasis is on understanding the application of a wide range of modern machine learning techniques to specific data analysis scenarios rather than mastering the techniques' theoretical underpinnings. The course covers methods aimed at supervised learning (e.g., prediction, classification) and unsupervised learning (e.g., clustering). This course is focused on applying machine learning with Python, and students will learn to develop and run Jupyter Notebooks and scripts. Nonetheless, students will gain an understanding of each algorithm, allowing them to assess performance and choose the best model for real-world datasets.

Learning Objectives

By the end of this class, students will be able to do the following:

1. Produce, comprehend, and run Python code for commonly used machine learning methods (supervised and unsupervised learning).
2. Apply feature engineering methods to different types of data to improve the performance of commonly used machine learning methods.
3. Choose appropriate model/s for datasets by evaluating their performance and the advantages and disadvantages of each.
4. Analyze and describe the societal impacts of machine learning methods implemented in real-world dataset/s while considering ethics, bias, and fairness.
5. Leverage data storytelling to report insights on machine learning model outputs.

Pre-requisites

- 90-800 Exploratory Data Analysis and Visualization with Python
- 90-819 Python Programming II
- A statistics course such as 90-707, 90-711, or 95-786

Auditing

This course demands a quick pace, active participation, and intricate collaboration with your peers. Only officially enrolled students will receive credit, and attendance will not be permitted for anyone else.

Learning Resources

I will post lectures and materials weekly for this class, and we will not follow a single textbook. The following references can be helpful as supplemental material for students:

- Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani
  An Introduction to Statistical Learning: with Applications in R- Second Edition
- Andreas C. Müller, Sarah Guido. Introduction to Machine Learning with Python

All the books above are available online through CMU's Library or O'Reilly (you all have free access to all their books with your CMU email).

Keep an eye on Canvas for assigned readings and materials.

Course Schedule

Date (Week)	Topics
Week 1 Jan 15th - Jan 21st	● Introduction to Machine Learning Foundations with Python ● Introduction to Modeling and ML Concepts
Week 2 Jan 22nd - Jan 28th	● Introduction to Regression Models ● Using Regression Models for Prediction
Week 3 Jan 29th - Feb 4th	● Regularization ● Ridge and Lasso
Week 4 Feb 5th - Jan 11th	● Model and Variable Selection ● Validation
Week 5 Feb 12th - Jan 18th	● Introduction to Classification ● Classification Models
Week 6 Feb 19th - Jan 25th	● Classification and Feature Engineering ● Thursday: EXAM 1 (Prediction)
Week 7 Feb 26th - March 3rd	● Evaluating Classification Models ● Multiclass Classification
March 4th - March 10th	**SPRING BREAK - NO CLASSES**
Week 8 March 11th - March 17th	● Tree-based methods ● Decision trees (classification + regression)
Week 9 March 18th - March 24th	● Ensemble Learning ● Random Forest, Boosting

Week 10 March 25th - March 31st	● Unsupervised Learning ● Clustering
Week 11 April 1st - April 7th	● Dimentionality Reduction ● Thursday: EXAM 2 (Classification)
Week 12 April 8th - April 14th	● Re-cap ● Bias, ethics and fairnes
Week 13 April 15th - April 21st	● Market Basket and Association Rules ● Advanced Topics
Week 14 April 22nd - April 28th	● Course Review and Wrap-Up ● Thursday: EXAM 3 (Unsupervised Learning)
Exam Week April 29th - May 5th	● Final Project Delivery ● Final Project Presentations (date to be determined by Heinz Final Exam Schedule)

***Note: This schedule is tentative and subject to changes. The schedule is only applicable for the current Spring 2024 cohort and may be changed as conditions merit it. If our pace is faster or slower than anticipated, then I will add or delete items at my discretion. Any changes will be announced in class (and on Canvas) and class announcements supersede the schedule presented here.*

The dates and modules on Canvas follow this same tentative schedule and will be updated on a weekly basis.

Course Assessments

The final course grade will be calculated using the following categories:

Assignment	Grade Percent
Final Project (multiple milestones throughout the semester)	36%
Homeworks (3 in total)	18%
In-class Challenges and Participation	16%
Exams (3 in total)	15%
Labs (12 in total)	11%
Out of-class work (exit tickets, surveys, etc.)	4%
Total Grade	100%

*All lectures, labs, and assignments are in the **Pittsburgh, PA timezone.***

In-person Attendance

To fully benefit from the material, it's essential to attend all classes and labs in person, as this class is designed to be highly interactive. For this reason, attendance during lectures will be assessed through in-class challenges and participation (see below).

In-class Challenges and Participation (16%)

During lecture time, in-class challenges will be given to students, which should be completed during class and submitted before leaving the classroom, regardless of how much is completed. These assignments are designed to help students practice and apply class content through small challenges. In the majority of these challenges, students will work with others in randomly assigned groups to complete the challenges, which will be evaluated based on effort and not on accuracy. Groups will have the opportunity to attempt and discuss these challenges and later share their findings with the rest of the class. This will provide everyone with the chance to learn from each other's diverse experiences and backgrounds. The correct answers to these challenges will be reviewed during class, and students will be able to access the solutions after they have been revised. There are in-class activities that will be individual and will be prompted as such.

Aside from the class challenges it is every student's responsibility to read and prepare for class with the materials provided on Canvas. As such, students are expected to participate in class. Class participation will be graded and noted by in-class engagement, including attempting in-class challenges, and asking relevant questions based on a conscious review of the required readings (and materials), lectures, and comments made by their peers. Class participation also involves being an active and contributing team member when assigned to group in-class challenges. The lack of attendance, and the use of mobile devices, including phones and laptops, will count against your participation grade.

Rubric for grading participation

	[Full credit]	[No credit]
[Deliverable]	Deliverable was submitted on time with thoughtful effort.	No deliverable was submitted.
[Participation]	Actively working together with group members and/or asking/answering questions during the class session AND not using mobile devices or laptops for non-class work.	No discussion, collaboration, or conversation with peers or instructors OR missing 10 or more minutes during the class session.

Clarifications on attendance:
- Classes begin on time and students are expected to be on time for every lecture (in their assigned section)
- During the semester every student has **two free absences**. Prof GS and the TAs do not need to know the details of the absence, you are all allowed two free absences. Once you have used your two free absences, your following absences will count against your grade for participation.
- Even if you are absent you can still receive half credit for submitting the in-class assignment individually within 24hrs.
- If you are experiencing an emergency (family, health, etc) that will last more than two consecutive absences please contact Prof GS (ggongora@cmu.edu) immediately.
- Please check the Late Attendance Policy below.

Late Attendance Policy

Obviously, the occasional emergency can disrupt even the most organized person's schedule; however, habitual lateness is considered unprofessional and inconsiderate. Students who arrive habitually late to class (more than ten minutes late) during the semester may have their grades lowered. If circumstances hinder your ability to arrive on time regularly, please speak with Prof GS (ggongora@cmu.edu) to discuss arrangements. This is consistent with Heinz College's norms and expectations. If a student arrives more than 10 minutes late to class they can still attend the class but they will be marked as absent for that lecture.

Homeworks (18%)

There will be a total of three homeworks, each counting for 6% of the final grade. Homeworks are individual assignments and will be graded on accuracy. Homeworks are expected to be coded in Jupyter notebooks, updated in your GitHub repository and submitted through Gradescope. The purpose of homeworks is to reinforce what students have been learning in class and in lab assignments. Lab assignments are done in a guided manner, while homeworks are done individually and will build and expand on lab assignments. Students will have two weeks to complete these coding homeworks at their coding pace. Homeworks are meant to challenge students; therefore, students should expect a higher complexity and time commitment than in lab assignments and in-class challenges, as well as expanding their skills and knowledge. It is advisable to start homeworks early, seek help early on, and revise past materials (lab assignments, in-class challenges, pre-work challenges) before starting the homework. Starting homework the day before the deadline will guarantee a poor performance on the assignment.

To account for "life happens" every student has one 24-hour pass to use at their discretion in the homeworks. This means that every student can apply this pass to waive the 24-hr late penalty on one of the homeworks. Students do not need to inform us of the late pass, the instructional staff will take note of late homeworks. On your first late homework, if you submit it within 24 hours you will still be eligible for 100% of the grade. If you submit 24-48hrs late on this first late assignment you will only be eligible for up to 80% of the original grade. After the 48hrs your homework will not be counted towards your grade. If you have a second or third late homework you may have up to 24 hours to turn but you will only be eligible for up to 80% of the original grade.

Exams (15%)

Throughout the semester, there will be a total of three exams, each worth 5% of your final grade. These exams will be done during class time (dates on Canvas) and in pen and paper, there is no coding involved. The exams will be in a case study format, and you will be exercising your ML problem-solving skills and reasoning to work through the exam individually. The exams are closed books, you are only allowed to bring a pen.

There are no make-up exams offered; therefore, if you miss an exam you will be forfeiting 5% of your final grade. Missing an exam because you overslept or had another commitment is not valid excuse. If you experience an emergency (.e.g. you have Covid, you are in the hospital, you have a death in the family, or similar), please contact Prof GS before the exam. Honesty and prompt communication go a long way in this course, and Prof GS will work with you through your situation.

Labs (11%)

Our weekly lab sessions will take place every Friday and will consist of purely coding exercises related to the topics seen that week. Lab sessions are an opportunity to learn and make mistakes subsequently preparing students and helping students perform better on their homework assignments. Lab sessions are also designed to reinforce the materials seen that week in lecture. The lab session is meant for students to work with others on the lab assignment and receive immediate feedback from the TAs - which is highly recommended - although attendance is optional. If students decide not to attend the lab session they must complete the lab assignment individually, and instructors will

not answer questions via Piazza or email about lab assignments. Lab assignments will be released at the beginning of the Lab session (Friday morning), and students have until the immediate Sunday at 11:59 PM to submit their completed lab assignments. Everyone is expected to complete the lab assignment, and they will be graded on completion. Students will be provided with the answers to lab assignments after the deadline. There is a total of 12 labs, but only 11 will count towards your final grade; this means that at the end of the semester, your lowest graded labs will be dropped, and only 11 labs will count towards your final 11% grade.

Out of Class Work (4%)

During this semester students will have a limited number of these assignments. These assignments are the easiest way to get points in this class and students should take advantage of that! The assignments for this category include: Syllabus Quiz, weekly feedback tickets (12 in total, starting on week two), a beginning-of-semester survey, and an end-of-semester survey. These assignments will be graded on completion.

Final Project (36%)

During this semester, teams of 2-3 students will work on a main project to create a cohesive ML project. The groups will be picked by the students, along with some requirements provided by Prof GS. This assignment is significant and time-consuming, as it tests various aspects of the course. To ensure that all teams receive feedback and have the opportunity to improve their projects gradually, the final project will be split and evaluated throughout 6 sprints and then a final delivery and presentation. The final project grade will be distributed as follows:

Assignments	Description	Weight (%)
Grade (Sprint 1-3)	You will receive a grade of your current progress (sprints 1-3) at the beginning of Week 8.	3%
Grade (Sprints 4-6)	You will receive a grade for these sprints(sprints 4-6) along with the grade for your Final Project.	5%
Final Project Delivery and Presentation	Delivered and presented during Final Exam week. Grades will be available before the deadline for final grades	28%
Total		36%

The sprints work in the following manner:
- For every sprint, you will receive feedback (not a grade) to ensure you are on track and you can incorporate that feedback into your final project.
- To give you a sense of where you stand in terms of the grading of your project during Week 8, you will receive a grade for your first three sprints. This grade will assess both the quality of your project (so far) and how well you have incorporated the feedback provided throughout these sprints.
- Sprints 4-6 will be graded along with your final project delivery. The grade for these sprints will be, again, dependent on the quality of your work during those three sprints and how well you have incorporated the feedback provided throughout these sprints.
- Please take a moment to review the dates and description of each sprint.

Sprint	Due Dates	Description
Sprint 1	Thursday, Feb 1st before 11:59 PM	Team assemble and contract
Sprint 2	Thursday, Feb 15th before 11:59 PM	Project proposal
Sprint 3	Thursday, Feb 29th before 11:59 PM	Data cleaning, initial visualizations, and explorations
Sprint 4	Thursday, March 21st before 11:59 PM	Initial ML models
Sprint 5	Tuesday, April 9th before 11:59 PM	Tunning Models
Spring 6	Friday, April 26th before 11:59 PM	Final project output and presentation preview

Students will present their final outcome during the exam week of this semester (Spring 2024). Please plan to be in Pittsburgh, as the presentations will be in person, and the final's schedule is dependent on Heinz' final exam schedule. The final exam schedule gets released by the University Registrar during the semester.

Final projects will be evaluated on their accuracy, thoroughness, and overall quality. Rest assured that Professor GS and the TAs are dedicated to supporting your success in this course and have confidence in your ability to produce exceptional work. Furthermore, ensure that your code is comprehensive and runs smoothly without any errors.

Final Course Grade

The final course grade will be calculated using the following scheme:

Grade	Letter
97 - 100 %	A+
94 - 96.99 %	A
90 - 93.99 %	A-
87 - 89.99 %	B+
83 - 86.99 %	B
80 - 82.99 %	B-
77 - 79.99 %	C+
74 - 76.99 %	C
70 - 73.99 %	C-
0 - 69.99 %	R

Late Work Policy

Due dates for every assignment will be provided and posted through Canvas. Unless otherwise stated, assignments are due on those days.

Type of Assignment	Policy Description	Accepted Late
Out of class work	These are the easiest of points to earn in this class and you will be given several days to complete them.	No
In-class Challenges and Participation	In-class challenges are opportunities for you to practice and get immediate feedback from your peers and the instructor. **For this reason, you may only earn credit for these challenges by attending class**. If you miss a class, keep in mind that there will be plenty of opportunities for you to gain points in this category during this semester. Please read "In-class Challenges and Participation" section for more details on absences.	No
Homeworks	In order to receive helpful and prompt feedback from Prof GS and TAs, it is important to submit your homework on time. However, if you encounter any unforeseen circumstances that prevent you from submitting your homework by the deadline, you may have up to 24 hours to turn it in. Please note that any submissions made during this 24-hour window will only be eligible for up to 80% of the original grade.	Yes - One 24hr free late pass. After that 24hr late penalty.
Labs	Lab assignments are designed to be completed during the allotted lab time, and students are encouraged to submit lab assignments at the end of the lab session. However, if you need additional time to complete your lab assignments, you may do so within the weekend, which means that everyone can submit the lab assignment before Sunday at 11:59 PM. Consequently, submitting later than this deadline will not be considered.	No
Final Projects	Final projects will be due during exam week, and in order for the instructional team to be able to grade them and submit final grades on time, final projects must be submitted by the deadline and will not receive a grade if turned in after the deadline. Sprints will also not be eligible for a grade if submitted late.	No

Re-Grades

If you wish to request a re-grade for homework or lab assignments, you must do so within three days of the grades being posted using the provided form provided on Canvas. To ensure clarity, please provide short and concise explanations in the form. Please note that full essays will not be accepted. When we receive your request, we reserve the right to review the entire assignment, which may result in an increase, decrease, or no change to your overall score.

However, we will not consider regrades if the score change is expected to be less than 5% unless there was an obvious error in tabulation or overlooked material. We will correct any tabulation errors, but it is not our responsibility to interpret your intentions or assumptions.

It is important to note that re-grade requests are not accepted for final projects, in-class challenges, or out of class work.

Programming and Computing Requirements

All coding assignments will be completed via Jupyter Notebooks. During the first week of classes, you will create a GitHub account, and every homework and lab will be delivered via GitHub. Jupyter Notebooks and the assignments can be completed on any type of computer and operating system. However, the settings you've chosen and the software you've installed might affect how Jupyter and Python run. I encourage you to follow our recommended settings (which you can find on Canvas) to ensure the smooth completion of your assignments and so we can better assist you when you encounter difficulties. Otherwise, we may not be able to help if you encounter any technical issues. Debugging and fixing computer problems is a skill that we want you to develop, and we will provide resources to help you learn how to do this independently. We suggest you keep track of these as you go through the course and in other courses in case you encounter problems.

Course Communication

Course Format

This is an in-person course held synchronously, meaning that students must attend all lectures in person. Lab attendance is highly encouraged. All lectures and labs will be held in person, and there will be no Zoom recording of them. Students are strongly encouraged to ask questions during class.

Office hours may be held through Zoom or in person. Please keep an eye on Canvas for the links and further information.

Laptops and Phones

This course involves regular use of technology during class, and at all times, laptops will be required. Research has shown that divided attention is detrimental to learning, so I encourage you to close any windows not directly related to what we are doing while you are in class and to close your laptops altogether unless needed in class. Please turn off your phone notifications and limit other likely sources of technology disruption, so you can fully engage with the material, each other, and me. This will create a better learning environment for everyone.

Recordings

If there are unforeseen circumstances and Prof GS needs to record the class, the recording will be available to all students. Recordings of class sessions are covered under the Family Educational Rights and Privacy Act (FERPA) and must not be shared with anyone outside your course section. The purpose of these recordings is so students in this course (and only students in this course) can watch or re-watch past class sessions.

No student may record any classroom activity without express written consent from Prof. GS.

Office Hours

Our team, including Prof GS and TAs, offers around 14 hours of office hours each week. We strongly urge you to utilize this opportunity to seek help and support whenever needed. If you are unable to attend in-person office hours, we have a Piazza platform where we can assist you in finding answers to your questions. We are dedicated to providing you with the best possible assistance.

How to Ask Questions

Your first point of contact in this course for asking questions should be Piazza. The Piazza forum should be used for general course-related questions that may interest others in the class. The quicker you begin asking questions on Piazza, the quicker you'll benefit from the collective knowledge of your classmates and instructors. We encourage you to ask questions when you're struggling to understand a concept.

In Piazza, we will have dedicated channels (e.g. Homework 1, Lectures, Final Projects) where you can post questions in a public or private setting. A public post will be seen by the entire class, while a private post will only be seen by Prof GS and the TAs.

Examples of questions that are suitable for a public post:
- Could you clarify __ point from Question 3)a) in the Homework?
- When is the final project due?
- What version of yellowbrick do I need to install if I'm getting the following error (insert screenshot)

Basically, any general, conceptual, or logistics question that does not involve your code or violates academic integrity, you can write it as a public post. Be sure to make these posts cohesive and comprehensive so that everyone can get the same understanding from your questions.

Examples of questions that are suitable for a private post:
- Several assignments are individual (without any collaboration); therefore, please make any specific questions (revealing code or your answer) as a private post (including Prof GS and the TAs).
- Please post private questions to Prof GS and all the TAs.

Piazza Threads and Follow-ups

Questions posted on Piazza should be treated as separate posts. If you have a follow-up on the same question/topic, then it can be added to the original post.

Please refrain from using a single Piazza question as a thread of questions or a conversation. For *every different question/topic, there should be a different Piazza post*. If you chain unrelated questions in a post, we will ask you to repost the question as a separate post before answering. This will allow for better clarity for everyone.

Monitoring Piazza

We will be monitoring Piazza every day to answer your question within the first 24 hrs during weekdays. Prof GS and TAs also need sleep and might not be awake to answer posts beyond 9 pm.

Bear in mind that answers to Piazza posts (and emails) will be within 48 hrs during the weekends; therefore, we encourage you to begin your assignment early and be proactive about asking questions during weekdays.

We encourage you to check Piazza before sending any questions via email to any TAs; a lot of times, the answer to your question has already been posted by someone else. In case the answer has already been posted, the TA will refer you back to Piazza. If you email new questions that are not of a personal nature (like grades, standing in class) your question will be directed to Piazza so that the answer can be useful for everyone. All students are invited to refer to the Etiquette section at the end of this syllabus.

Every student in the course will be enrolled in the Piazza. If you have not been enrolled, please email Prof GS (ggongora@cmu.edu). If you have any problems or feedback for the developers of Piazza, please send an email to mailto:team@piazza.com.

Canvas

Assignments and class information will be posted on Canvas. Take note of the following sections on Canvas:
- Home - Location, credits, office hours, and general information about the course.
- Assignments - this tab contains all the assignments posted for this course.
- Announcements - Make sure you receive our Canvas Announcements, as we will be making all announcements through there and double posting them on Piazza.
- Modules - this tab will have all the materials for the lectures, labs, and any important dates (e.g. no classes, assignment deadlines, etc) posted. The modules are arranged by week. There are extra modules with materials and with answers to Labs and Homeworks. The solutions to class exercises will be posted in the weekly module they correspond to.
- Piazza - short-cut tab to Piazza.

Prof GS and the TAs do not use Canvas Inbox messages, therefore, refrain from sending any communication through that feature.

Email Communications

Prof. GS recognizes that sometimes "life happens." For this reason, if you have any issues during the course that interfere with your performance in the class (e.g. health, life-related, personal, etc.), please reach out to Prof. GS through email (ggongora@cmu.edu). Honesty, communication, and transparency are highly welcomed and encouraged in this course.

When emailing me, please place in the subject line the course number/section and the topic of the email (i.e. 90-800 – Student name). This will help me tremendously in locating your emails when I scan the hundreds of emails that seem to make it into my box each day.

Please be advised that sending an email to your instructors does not create the responsibility or obligation to respond to it. Sending us an email does not shift any responsibility from you to us. You are still responsible for the on-time, high-quality completion of assignments and projects. For emails that are non-emergencies, you should receive a response within

To protect your academic privacy, emails will not be used to send grade reports either, grades are posted on an individual basis via Canvas.

Academic Honesty and Integrity

All assignments - unless stated - are to be done individually. You are encouraged to discuss only lab problems and certain in-class challenges with your fellow students. However, the work you submit for the Lab must be your own. You must acknowledge in your submission any help received on your assignments. That is, you must include a comment in your lab submission that clearly states the name of the student, book, or online reference from which you received assistance. Again, this is perfectly acceptable and will not be penalized so long as your work is your own. Submissions that fail to acknowledge help from other students or non-class sources properly will receive no credit. Copied work will receive no credit. Any and all violations will be reported to the Heinz College administration.

All students are expected to comply with the CMU policy on academic integrity. This policy can be found online at http://www.cmu.edu/academic-integrity/. Every line of text and line of code that you submit must be written by you personally. You may not refer to another student's code or a "common set of code" while writing your own code. You may copy/modify lines of code that you saw in the lecture or lab, but you must also reference and acknowledge the source.

While Prof GS encourages you to be helpful to your classmates, you must understand that the work you turn in for evaluation or credit must be your own. You are welcome to discuss general course content, requirements, and technology issues with other students. You are not welcome to offer, or ask for substantial material assistance to or from other students in completing specific aspects of graded assignments for individual credit. If you have any doubt about a particular situation, ask yourself this question: "How would I feel if I observed another student or students engaging in this particular behavior?"

Here are some examples of academic integrity violations related to code:
- Copying line-by-line (or substantial amounts) of another student's code (classmate or peer).
- Showing your code or working with another student to complete individual assignments (e.g. Homeworks)
- Copying line-by-line of online (or substantial amounts) code (e.g., Stackoverflow, Chat-GPT) or code found in one of our textbooks.
- Posting class materials or assignments online to get help. You can only post questions on our Piazza.
- Posting class materials or assignments in a public repository, therefore allowing other classmates to replicate your code.

Other types of academic integrity violations:
- Taking another student's name tag when they are not physically present.
- Writing your name on an in-class assignment when you did not attend that class in person.
- Writing someone else's name on an in-class assignment when they did not attend that class in person.

Note from Prof. GS: "viewing someone's code, sharing code with another, or giving specific instructions, verbally or in writing, for work in any of the phases is definitely out of bounds. Do not post your code to a public repository where it is available to all or a private repository where it is viewable by other students." If there is any doubt whether it is allowed, please see Prof. GS or our TAs to get clarity on the matter. Better safe than sorry!

Any student who turns in work for credit that is identical or similar beyond coincidence to that of another student may face appropriate disciplinary action at either the college or university level. Your reputation among your peers and among the CMU faculty is one of your most valuable assets. Do not risk damage to your reputation or academic career by engaging in behavior that could be interpreted as dishonest or unethical.

To be clear, any academic integrity violation will result in a zero in that assignment and you will be immediately reported to our Associate Deans and face appropriate disciplinary action at the department, college, and even University level. Depending on the type of assignment you may also receive an additional penalty of two letter grades off your final grade in the course (e.g. if you were going to get an A, your final grade will be reduced to a C).

Use of generative AI (e.g., ChatGPT)

You may use generative AI programs like ChatGPT during the brainstorming and idea-generation phase for assignments. However, doing so cannot be considered a substitute for traditional research. Generative AI programs rely on predictive models to generate content that may appear correct but has been shown sometimes to be incomplete, inaccurate, taken without attribution from other sources, and/or biased. Any information generated by an

AI program should be cited like any other reference material. You are ultimately responsible for the content of the information you submit. However, you may not attempt to pass off any work generated by an AI program as your own.

Expectation Guidelines

Prof GS wishes to foster a safe learning environment where all students have the opportunity to learn and whose ideas and contributions are valued. All opinions and experiences, no matter how different or controversial they may be perceived, must be respected in the tolerant spirit of academic discourse. You are encouraged to comment, question, or critique an idea, but you are not to attack an individual. Our differences, some of which are outlined in the University's inclusion statement below, will add richness to this learning experience. Please consider that sarcasm and humor can be misconstrued in online interactions and generate unintended disruptions. Working as a community of learners, we can build a polite and respectful course ambiance.

Please read the guidelines for this course:
- Be respectful of your instructors, teaching assistants, and fellow students.
- We all communicate differently, and it's invaluable to learn from each other's backgrounds and experiences. Try to leave room for other students to join discussions in our group assignments and in class.
- There is zero tolerance for inappropriate language, as well as racist, sexist, or discriminatory language, in any context during the course.
- Additionally, any assignment, documentation, code, or files that contain inappropriate language, comments, or inappropriate references will incur an automatic 30% penalty without exemption.
- Be cautious in using the Internet language when communicating via Piazza, Email or Canvas. For example, do not capitalize all letters since this suggests shouting.
- Consider that there are different cultures and nationalities in this course; therefore using any vernacular and/or slang language could lead to a lack of understanding or misinterpretation.
- Keep an "open mind" and be willing to express your opinion in a manner that is respectful and still adheres to these guidelines.
- Do not hesitate to ask for feedback, we are all here to learn.
- We intend to start and end each class on time, so please be respectful to everyone by being on time and staying the entire session.

If you disagree with any guidelines or would like to add more, please contact me, I welcome your input as we craft a respectful environment in this course.

This course will actively engage in learning from how to build ML models to interpreting and assessing ML models. The course content is designed to be applicable to other courses, and I (Prof GS) hope you find it resourceful, insightful, and challenging. Embrace the challenges that come with the course as we apply our skills to different examples of healthcare, public policy, and social issues. To ensure effective learning, I encourage you to prepare with the recommended materials and readings, revise past materials, and keep in mind that coding has multiple solutions. We will work together to help you enhance your skills and apply them. By nature, ML is a repetitive process, where the answer will never be straightforward and will require you to keep iterating, trying avenues, and formulating a decision afterward. You will not be able to grasp these complex topics by only attending lectures; these concepts will require you to code constantly and build on your knowledge and skills repeatedly. You will benefit from practicing in your own time and spending time self-learning. This is an iterative process where you will make mistakes - as we all do - and learn from your peers, your instructor, your TAs, and your own mistakes.

If you have any questions about the course or coursework, please do not hesitate to ask. Let's have a great semester!

Research to Improve the Course

For this class, I am conducting research on student outcomes. This research will involve your work in this course. You will not be asked to do anything above and beyond the normal learning activities and assignments that are part of this course. You are free not to participate in this research, and your participation will have no influence on your grade for this course or your academic career at CMU. If you do not wish to participate or if you are under 18 years of age, please send an email to Chad Hershock (hershock@andrew.cmu.edu) with your name and course number. Participants will not receive any compensation. The data collected as part of this research may include student grades. All analyses of data from participants' coursework will be conducted after the course is over and final grades are submitted. The Eberly Center may provide support on this research project regarding data analysis and interpretation. The Eberly Center for Teaching Excellence & Educational Innovation is located on the CMU-Pittsburgh Campus and its mission is to support the professional development of all CMU instructors regarding teaching and learning. To minimize the risk of breach of confidentiality, the Eberly Center will never have access to data from this course containing your personal identifiers. All data will be analyzed in de-identified form and presented in the aggregate, without any personal identifiers. If you have questions pertaining to your rights as a research participant, or to report concerns to this study, please contact Chad Hershock (hershock@andrew.cmu.edu).

Student Wellness

Disability Resources

If you have a disability and need special accommodations in this class, please review the steps listed by the [Office of Disability Resources](#), I encourage you to contact them at access@andrew.cmu.edu or contact Catherine Getchell, Director of Disability Resources, 412-268-6121.

If you already have an accommodations letter from the Disability Resources Office, I encourage you to discuss your accommodations and needs with me as early in the semester as possible. I will work with you to ensure that accommodations are provided as appropriate.

Diversity Statement

We must treat every individual with respect. We are diverse in many ways, and this diversity is fundamental to building and maintaining an equitable and inclusive campus community. Diversity can refer to multiple ways that we identify ourselves, including but not limited to race, color, national origin, language, sex, disability, age, sexual orientation, gender identity, religion, creed, ancestry, belief, veteran status, or genetic information. Each of these diverse identities, along with many others not mentioned here, shape the perspectives our students, faculty, and staff bring to our campus. We at CMU, will work to promote diversity, equity, and inclusion not only because diversity fuels excellence and innovation but because we want to pursue justice. We acknowledge our imperfections while we also fully commit to the work, inside and outside of our classrooms, of building and sustaining a campus community that increasingly embraces these core values.

Resources for Diversity and Inclusion:
- [Center for Diversity and Inclusion](#)
- [Intercultural Communication Center](#)
- [Office of Title IX Initiatives](#)

Wellness Statement

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. 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