Machine Learning

Spring 2021 | Syllabus | CIS 471-01

Class Information

Professor
Lauren Williams, PhD

Meeting Times
MW 3:30 - 4:45

Meeting Location
Library 126

Zoom Office Hours

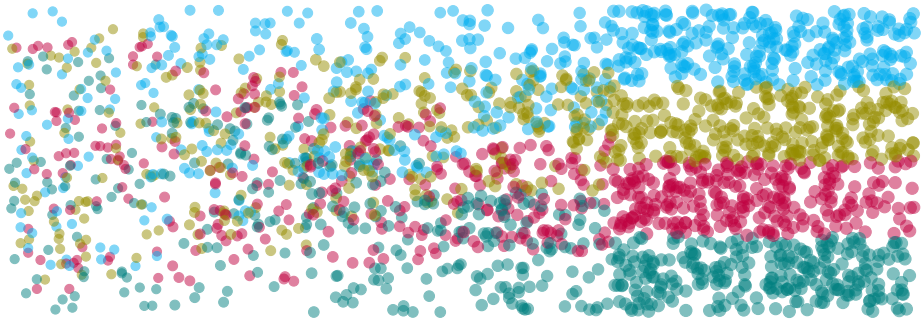
Monday 12 - 12:50
Tuesday 9 - 10, 12 - 1:50
Wednesday 12 - 12:50
Thursday 9 - 10

Contact

Email
lwilliams@mercyhurst.edu

Office
Old Main 404

Office Phone
(814) 824-2226



Course Description

This course will be an introduction to the methods, algorithms, and best practices of machine learning. Topics will include a review of essential mathematics and statistics, supervised vs unsupervised learning, and selected methods of classification, regression, clustering, and forecasting. We will also see a brief introduction to advanced concepts in machine learning, including neural networks and deep learning. Several of these algorithms will be compared and implemented using Python and many of its popular machine learning libraries.

Course Objectives

Upon successful completion of this course, a student will:

- be familiar with the foundational mathematics and statistics required to implement and interpret a machine learning solution.
- understand the basic categories of machine learning, such as supervised and unsupervised learning.
- determine which methods of machine learning, if any, is best suited to solve a particular problem.
- recognize the limitations and potential drawbacks of machine learning.
- have a working knowledge of several popular machine learning libraries from the Python programming ecosystem.

Required and Suggested Materials

Software and Tools

Machine learning can be done in just about any high level programming language, with Python and R among the more common in use today. Any demonstrations of machine learning techniques for this class will be done using Python, including the popular library scikit-learn.

You are free and encouraged to use any editor and compiler you prefer. I will be using the Anaconda platform, which is available for free for Windows, Mac, and Linux.

Books

A textbook will not be required for this course, but when you're looking for more information, here are some highly recommended sources. Many of our class lectures will draw material from these sources:

- **Python Machine Learning**, by Sebastian Raschka and Vahid Mirjalili, Packt Publishing (3rd Edition)
  Link to Purchase
  From the publisher: "Python Machine Learning, Third Edition is a comprehensive guide to machine learning and deep learning with Python. It acts as both a step-by-step tutorial, and a reference you'll keep coming back to as you build your machine learning systems."
- **Introduction to Machine Learning with Python**, by Andreas C. Müller and Sarah Guido, O'Reilly
  Link to Purchase
  From the publisher: "You'll learn the steps necessary to create a successful machine-learning application with Python and the scikit-learn library. Authors Andreas Müller and Sarah Guido focus on the practical aspects of using machine learning algorithms, rather than the math behind them. Familiarity with the NumPy and matplotlib libraries will help you get even more from this book."
- Various Machine Learning Titles by Jason Brownlee
  Link to Purchase
  Nearly 20 textbooks available as PDFs, covering beginning, intermediate, and advanced machine learning topics as well as prerequisites such as statistics and linear algebra.

There are many textbooks available for machine learning, and there tends to be a great deal of variety with regards to approach (theoretical vs applied), topics covered, programming ability, and required background information. There are even more websites devoted to topics in machine learning. All of this means that investing in a textbook is challenging, as it's hard to pick the right one and may not be worth the risk given the vast quantity of free, high quality content available elsewhere.

Grading

Assignments 60%

Assignments will be posted regularly throughout the semester. Some will be short assignments that can be completed within Blackboard, and are meant to ensure you are keeping up to date with the material as it is covered in class. Other assignments will require you to submit code or the results of code, using techniques similar to those demonstrated in class. Some assignments will be worth more than others towards your final grade. No assignment grades are dropped. Late assignments will be accepted for up to two weeks after they are due with a reduction in the maximum possible score.

Exams 40%

We will have two exams during the semester, both equal in weight. Exams will focus more on concepts, definitions, and techniques of machine learning, as opposed to a traditional programming exam. Exams will be delivered on the dates stated in the course syllabus. Please let me know as soon as possible if you will not be able to take an exam

Grade Scale

F	D	C	C+	B	B+	A
0-59	60-69	70-78	79-82	83-89	90-92	93-100

Other Policies

Missed Class
You do not need to let me know if you'll have to miss class. Just make sure you keep up to date with the schedule, and take advantage of the Zoom recordings of class if you are unable to attend in person.

Missed Exam
If you know in advance that you will not be able to take an exam on the day it is scheduled, please discuss with me to make arrangements for an alternative time. Make up exams will not be given for undocumented reasons (such as illness or university related travel) unless arrangements are made in advance.

Email
I will attempt to respond to emails promptly. For your own protection, please use your Mercyhurst email address. Emails received after 6 pm may not receive a reply until the next business day. If you have not received a response from me in over 24 hours on a weekday, feel free to send another - I may have missed your first message.

Extra Credit
Out of fairness to all students in the class, I do not give extra credit opportunities to improve your final grade.

Academic Honesty
All students in this class are expected to maintain a high standard of academic integrity. Any instance of plagiarism or cheating will result in a 0 on the assignment or exam. A second incident will result in a submission of an academic dishonesty report to the university, and may result in an F in the course.

Semester Schedule

Week	Topic
1 Jan 25 / Jan 27	**Course Introduction and Statistics** Course overview, a look at exactly what machine learning is, and a brief introduction to some concepts from statistics we'll need to understand our models.
2 Feb 1 / Feb 3	**Linear Algebra** A crash course in linear algebra, a field of mathematics that forms the basis of many of the algorithms and approaches we'll use in machine learning. Vectors, matrices, matrix operations, inverses, eigenvalues, eigenvectors, and matrix factorization will be summarized.
3 Feb 8 / Feb 10	**Probability, Graph Theory, and More Mathematics Tools** We'll continue our review of essential mathematics for machine learning with an overview of counting methods and probability, calculus, and the basics of graph theory.
4 Feb 15	**Python and Related Tools** Before we start building our models, we'll need to know about some Python libraries that are of interest to data scientists and students of machine learning. (No class Wednesday of this week)
5 Feb 22 / Feb 24	**Errors and Model Evaluation** Machine learning is not an exact science. We'll look at errors and bias introduced by our training sets, ethical considerations, the problem of overfitting, and how to detect issues within our models.
6 Mar 1 / Mar 3	**Regression** Regression involves building a model used to predict numerical output. We'll look at linear and logistic regression, along with ways to measure accuracy and improve fit with regularization.
7 Mar 8 / Mar 10	**K-Nearest Neighbors** This algorithm can be used for classification and regression problems. It's very simple to implement and understand, but also problematic with respect to efficiency.
8 Mar 15 / Mar 17	**Midterm Exam I** Monday's class will be a review of the material we've seen so far and tips for the first exam, which will be delivered on Wednesday.
9 Mar 22 / Mar 24	**Decision Trees** Decision trees are natural to understand but difficult to construct, as we have to balance accuracy and efficiency. We'll see how to use decision trees in classification and regression problems.
10 Mar 29 / Mar 31	**Support Vector Machines** Support vector machines refer to models used for classification and analysis that rely heavily on mathematics to implement. We'll look at cases where the data is linearly separable, and how to choose and implement kernel functions to improve classification of non-separable data.
11 Apr 5	**Recommendation Systems** Systems that recommend products, movies, music, and even friends are increasingly common. We'll look at the general ideas behind these systems and what approaches are taken to create a fast but reliable recommendation application. (No class Wednesday of this week)
12 Apr 12 / Apr 14	**K-Means Clustering** A useful and simply unsupervised learning method used primarily for classification and early data analysis.
13 Apr 19 / Apr 21	**Time Series and Forecasting** Machine learning involving data with a time attribute, such as seasonal changes, require special consideration. We'll look at common approaches and problems with time series analysis.
14 Apr 26 / Apr 28	**Deep Learning and Neural Networks** While both of these topics are outside the scope of an introductory course on machine learning, we should be familiar with the power and risks of using these models.
15 May 3 / May 5	**Midterm Exam II** Monday's class will be a review of the material we've seen so far and tips for the second exam, which will be delivered on Wednesday.

University Resources and Policies

ADA Accommodations/Academic Support

Mercyhurst University is committed to making reasonable accommodations for qualified students, and employees with disabilities as required by law. Please refer to the HUB https://lakersmercyhurst.sharepoint.com/sites/StudentsHub and select the Services tab, then ADA Accommodations from the dropdown for instructions to request an accommodation. You may also contact Susan Reddinger, ADA Coordinator, ADA@mercyhurst.edu, 814-824-2362, Egan Hall 200.

For students with questions about Academic Support, please refer to the HUB https://lakersmercyhurst.sharepoint.com/sites/StudentsHub and select the Academic Resources tab, then Academic Support for more information.

TITLE IX Sexual Misconduct/Sexual Harassment Reporting

Mercyhurst is committed to providing an environment free from sex discrimination, including sexual harassment and sexual violence. Please refer to the HUB https://lakersmercyhurst.sharepoint.com/sites/StudentsHub and select the Resources tab, then Title IX – Sexual Respect from the dropdown for more information. If you would like to file a sexual misconduct complaint, please contact Dr. Laura Zirkle, Interim Title IX Coordinator and VP for Student Life, titleix@mercyhurst.edu, 814-824-2362, Egan Hall 314. Please be aware that in compliance with Title IX, educators must report incidents of sexual assault/harassment, stalking, and domestic/dating violence. If you disclose any of these situations in class, in papers, or to me personally, I am required to report it to the Title IX Coordinator (or any of the Deputy Title IX Coordinators).

Academic Honesty

Students are required to uphold academic integrity throughout the course. In particular, the use of unauthorized materials or collaboration on assignments or exams and other incidences of academic dishonesty will be handled according to the policies set forth in the Student Handbook.

Course Evaluations

Near the end of the semester, you will be asked to complete an online course evaluation. The evaluation will be completed in class during the last two weeks of the semester using any laptop, tablet, or mobile device. The response tool allows you to note aspects of the course that helped you learn, as well as aspects that might be modified to help future students learn more effectively. You will receive an email letting you know when the evaluation window for our class is open. Please note that these course evaluations are anonymous and instructors do not see the results until after the grades for the course are submitted.

COVID-19 Information

This is sure to be an unprecedented semester! While we cannot know what the next few months will bring, we must all work together to keep our campus community safe and healthy.

This page features information regarding policies of the University (in italics) as well as comments, suggestions, and requests that pertain to our class specifically.

A/B Meeting Schedule

Currently, our class is small enough that we will not need to break into A/B groups. We can all fit safety and socially distanced in our classroom, so please plan to attend if you do not have a designation of "approved for remote study".

Face Masks

As per the COVID-19 Prevention, Mitigation, and Response Policy, Mercyhurst University is requiring that all members of the campus community wear a cloth or disposable face covering over their nose and mouth when on campus. Please refer to the policy for specific details as to where and when face coverings are required. Students may use their own face coverings or those provided by the University. A student in need of a face covering should email covid19@mercyhurst.edu or call 814-824-3600 to find the nearest location where face coverings are available. The University's Mask/Face Coverings Policy will be enforced in this class.

Sanitation and Safety

In keeping with the COVID-19 Prevention, Mitigation, and Response Policy, students are expected to use hand sanitizer and to wipe down their desks using disinfectant wipes when they enter and exit the classroom. Classrooms have been provided with sanitizer and disinfectant wipes for student and faculty use.

Eating and Drinking in the Classroom

In light of the COVID-19 situation, Eating is not permitted in classrooms, labs, or other academic spaces. A water bottle or cup with a lid, and straw preferably, is permitted to be used in classrooms and labs to help prevent a student from becoming uncomfortably parched. Masks should be pulled only slightly away from the bottom of the face to take a quick drink and immediately replaced to covering the mouth and nose.

Class Dismissal and Congestion Prevention

In keeping with the COVID-19 Prevention, Mitigation, and Response Policy, faculty members and students should take steps to avoid crowding outside of classrooms, in hallways, and any enclosed area in university buildings. All rooms will be designated with signs indicating maximum capacity for specific instructional use. These must always be adhered to. Students waiting to enter classrooms or exiting classrooms should always maintain a minimum of 6 feet of distance from others. Class time endings may be adjusted when necessary to minimize overcrowding or congestion.

COVID-19 Information, Continued

Seating Chart

In compliance with federal and state regulations, the University must be able to conduct contact tracing if there is a positive test or an outbreak; therefore, seating charts are mandatory for all in-person classes. Students will be required to sit in the same seat in the classroom each time they attend class. The seating chart will be available for review for purposes of contact tracing.

Paper Sharing Policy

We will not be exchanging paper this semester. Supplemental materials will be distributed and made available electronically. Assignments and exams will be submitted electronically as well. You are welcome to bring your own paper to class to take notes, but you may not pass paper to a classmate or to me.

Attendance and Missed Classes

Attendance at all classes is expected. However, it is important that students and course instructors adhere to the university's COVID-19 mitigation policies and strategies. As such, a student who misses class due to illness or suspected illness within the context of those policies will not be penalized and will be provided sufficient means to make up any missed course content or work and remain actively engaged in the class.

The word "attendance" has a broader definition than usual this semester. While attending class is preferred, please do not feel obligated to come if you are feeling ill. You are free to join and participate in the live Zoom meeting in lieu of attending a class meeting, regardless of your A/B schedule. If you're not up to joining in, please watch the recording of the class meeting and the associated video lectures when you're able to.

If you are unable to attend class (or join the live Zoom meeting) for more than a few days, please let me know as soon as possible. I am happy to work with you in building a plan that allows you the time off you need without risking your academic progress.

Potential Class Changes

It is my hope that we remain able to meet in person as scheduled for the entire semester. However, there is a very good chance that our plans will change, and without much notice.

If we are unable to continue meeting in person, you will still have the opportunity to "attend" class via Zoom during our regularly scheduled times. The course content on Blackboard was created to help fill in any unexpected gaps in meetings or attendance.

Our highest priority is to remain healthy and safe. We will all need to remain responsible, flexible, and understanding to make this semester a success, and I have full confidence that we will be able to achieve that goal.