CISC 5800 Machine Learning
Department of Computer and Information Science
Dr. Daniel D. Leeds, Fall 2020

Class times: Thursday, 5:30 – 7:45pm
Location: Usually online, sometimes in person LL 311
Instructor: Prof. Daniel D. Leeds
Office: Usually online, sometimes in person at location TBA
E-mail: dleeds@fordham.edu
Office hours: Online Tuesday 4-5pm ; Sometimes in person Thursday 4-5pm

Course website: http://storm.cis.fordham.edu/leeds/cisc5800/
Also see Blackboard

Hybrid format for our class:
Most weeks, lectures and office hours will be online. They will take place live at the times specified above (Thursday evening, NYC time). These online lectures will often last roughly 2 hours. Sometimes live lecture will be shorter, with short asynchronous videos to be watched before or after the live session.
On some weeks, we will have a shorter in-person option for lecture and office hours, as well as a shorter live remote option at a different time. For example, we may have a live in-person lecture 5:30-6:30pm Thursday and a live online lecture 6:45-7:45pm. On these weeks, there often will be additional content recorded in asynchronous online videos you also will need to watch. In-person lecture days will be confirmed to the class at least one week ahead of time.

Texts: No text is required, but it will be useful to have one for reference.
"Pattern Recognition and Machine Learning", C. Bishop, 2007.
Bishop is a classic text in the field. I recommend Bishop for most students.

"Machine Learning: A Probabilistic Perspective", K.P. Murphy, 2013.
Murphy is a newer text. It provides more in-depth analyses we will not cover in class, but may be of interest to mathematically advanced students.

Our course site also links to course notes from Andrew Ng's Machine Learning course notes at Stanford University, which are particularly helpful.

Course description: Machine learning discovers and responds to patterns in the rich data sets of the world. It provides powerful tools for diverse fields, from software development to product marketing to scientific research. This course introduces a collection of prominent machine learning models. We will study both theory and implementation. Topics will include: Bayesian statistics, learning theory, support vector machines, dimensionality reduction, and graphical models.

Objectives: To understand the mathematical principles and algorithmic mechanics behind popular methods in machine learning. A student who successfully completes this course will be able to:

- Implement "basic" learning algorithms in the Python programming environment
- Apply diverse learning algorithms to complex data
- Understand the advantages and limitations of diverse learning approaches

Software: We will complete our programming assignments in Python.

- In Python, you can use numpy, scipy and matplotlib packages to write acceptable code for this class in Python. (You also may used the pandas package; you **may not** use packages like tensorflow, torch, or sklearn for this class.)

Attendance and class participation: As we do not directly-follow any one textbook and, instead, rely heavily on lectures and lecture slides, it is important to attend every class, and to arrive on time. One unexcused/unexplained absence is permitted for the semester. Attendance will be taken regularly. Please *actively* participate in class since this will make the course more interesting for everyone! Ask questions if you are unsure about something. If live attendance is impossible for you (including live online attendance), please let me know at the start of the semester and we will come up with a plan. We need to ensure every student has regular occasions to interact with the instructor and with the class.

Course assignments: There will be 4-5 homeworks and a final project assigned for the course. The homeworks usually will be announced at least a week before they are due, e.g., a homework announced on Thursday may be due the following Thursday. All assignments must be turned in on time through Blackboard.

Academic honesty: All work submitted in this course must be your own. You may discuss the assignment problems with other students generally, but you may not provide complete solutions to one another. Copying of any part of an assignment is never acceptable and will be considered a violation of Fordham's academic integrity policy. Violations of this policy will be handled in accordance with university policy which can include automatic failure of the assignment and/or failure of the course. See Fordham's Graduate Policy on Academic Integrity for more information.

In-class exams: There will two quizzes, one mid-term exam in October and a final in December. Quizzes will be announced at least a week in advance. The exact dates of the midterm and final will be announced at least 3 weeks in advance of the exams. All exams will be held online live.

Timing conflicts: If you have a significant issue and cannot complete an assignment on time, or cannot attend class on a certain day, whenever feasible let me know beforehand -- I tend to be reasonable in such cases. Examples of significant issues include personal illness or a religious holiday (give me at least a week's notice) on an announced exam day. In general, let me know of any significant issues that affect your performance early on.

Grading: The percentages given below are guidelines for both the student and instructor and may be changed as needed to reflect circumstances in the course. Any changes that occur during the semester will be minor.

Participation	5%	Mid-term	25%
Homeworks	20%	Final exam	25%
Final project	15%	Quizzes	10%

Tentative schedule

Schedule is subject to change as the class progresses. The last few days are scheduled as “catch-up” days as it is likely we will run behind at some point as the semester progresses. If we are **ahead** of schedule, I will add some additional topics.

Days with an in-person option will be officially confirmed one week ahead. Here I provide my best guess to help you plan.

August 27	Background: Probability, calculus, classifier review	
September 3	Bayes classifier	HW 0 due
September 10	Logistic classifier **– EXPECT IN PERSON OPTION**	
September 17	Logistic classifier / Support vector machines	HW 1 due
September 24	Support vector machines	Quiz 1
October 1	Dimensionality reduction **– EXPECT IN PERSON OPTION**	HW 2 due
October 8	Neural networks **– EXPECT IN PERSON OPTION**	
October 15	Midterm	
October 22	Neural networks	
October 29	Hidden markov models **– EXPECT IN PERSON OPTION**	HW 3 due
November 5	Expectation maximization	Quiz 2
November 12	Mixture models **–EXPECT IN PERSON OPTION**	HW 4 due
November 19	Learning Theory	
December 3	Convolutional Neural Networks	Final project due
December 10	Final	