Achim Koh, Visiting Assistant Professor
Pratt Institute School of Information
144 W. 14th St., 6th Floor
New York, NY 10011-7301
Email: akoh7@pratt.edu
URL: https://scalarvectortensor.net

INFO 697-03: Machine Learning
Fall 2019

Class Hours: Tuesdays, 6:30p – 9:20p
Office Hours: By appointment
Credits: 3
Prerequisites: INFO 654 Information Technologies, or by permission
Location: PMC 606

Course Description:

Machine learning is a rapidly growing field that develops algorithms for tasks such as data classification and prediction. These algorithms are programmed to operate and adjust themselves independently of human intervention (i.e., to learn), allowing data work to occur quickly and at scale. Machine learning is a key technology behind the automation across many social areas today, often branded AI.

This course offers an introduction to machine learning as a practical tool that we can use, and as a technological field with social implications. We will learn about key concepts in machine learning; survey a few key machine learning techniques, such as supervised methods for machine learning (regression and classification), which attempt to map data onto desired outputs, and unsupervised methods (clustering and association), which attempt to find structure within data itself; use openly available tools to implement these techniques on text and image data; learn how to assess the effectiveness of different techniques on particular datasets; and discuss basic issues that confront all machine learning methods.

Readings, class discussions, and hands-on sessions will be complemented by guest lectures (TBD) from machine learning practitioners. Students will be assessed via a final project developed throughout the course, in addition to the project proposal, presentation, class participation, and lab assignments.

This syllabus is a living document; expect it to evolve over the course of the semester. All changes will be communicated in class and the updated syllabus will be uploaded on LMS. Since this is a new course, your participation and input will be crucial in shaping it to your needs. Feel free to ask questions and give feedback or suggestions, in person or via email, as we move into the semester.

Course Goals:

The goals of this course are to:

- introduce students to key concepts and some common techniques in machine learning, as well as openly available tools
- help students to develop technical and critical thinking skills regarding machine learning
- enable students to conduct a machine learning experiment and communicate the result of their project

Student Learning Outcomes:

By the end of the course, students will be able to:

- describe different machine learning methods, including their limitations
- select an appropriate machine learning method for a given use case
- implement machine learning algorithms and assess their performance
- execute a machine learning experiment using openly available tools
- support the design of their experiment by discussing both the technical aspect and the topic's significance

---

Course Schedule and Readings:

> Week 1 - 8/27: Introduction
Lab: Getting started with Python

> Week 2 - 9/3: Machine learning, data, programming
Readings:

- Meredith Broussard, *Artificial Unintelligence*, ch.2-3 (13-39) – via LMS
- Liza Daly, "AI Literacy: The basics of machine learning" https://worldwritable.com/ai-literacy-the-basics-of-machine-learning-2e20f93e34b4
- Siddhartha Mukherjee, "AI Versus MD" http://web.archive.org/web/20170427141526/http://www.newyorker.com/magazine/2017/04/03/ai-versus-md
- Gideon Lewis-Kraus, "The Great A.I. Awakening" http://web.archive.org/web/20161215073155/https://www.nytimes.com/2016/12/14/magazine/the-great-ai-awakening.html

Supplemental Material:

- Paul Ford, "What is Code?" https://www.bloomberg.com/graphics/2015-paul-ford-what-is-code/

Lab: Working with data in Python

> Week 3 - 9/10: Classification #1
Due: Lab assignment #1 (submit by 9/9)

Readings:

- Broussard, *Artificial Unintelligence*, ch.7 (87-119) – via LMS
- Stephanie Yee and Tony Chu, A visual introduction to machine learning http://www.r2d3.us/visual-intro-to-machine-learning-part-1/

Supplemental Material:

- Gene Kogan and Francis Tseng, "Fundamentals, introduction to machine learning" https://github.com/ml4a/ml4a-guides/blob/master/notebooks/fundamentals.ipynb

Lab: Introduction to scikit-learn; classifiers

> Week 4 - 9/17: Classification #2
Readings:

- Johanna Drucker, "Humanities Approaches to Graphical Display" http://www.digitalhumanities.org/dhq/vol/5/1/000091/000091.html
- Lisa Gitelman (ed.), *"Raw Data" Is an Oxymoron*, introduction (1-14) http://raley.english.ucsb.edu/wp-content/Engl800/RawData-excerpts.pdf
- Mimi Onuoha, *On Missing Datasets* https://github.com/MimiOnuoha/missing-datasets

Lab: Classifiers continued; textual and image data

> Week 5 - 9/24: Models and abstraction; features and evaluation
Due: Lab assignment #2 (submit by 9/23)

Readings:
- Cathy O'Neil, *Weapons of Math Destruction*, introduction-ch.1 (1-31) – via LMS
- Alex Galloway, "Are Algorithms Biased?" http://cultureandcommunication.org/galloway/are-algorithms-biased
- Os Keyes, "Counting the Countless" https://reallifemag.com/counting-the-countless/

Supplemental Material:
- Julia Angwin et al., "Machine Bias" https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing
- Shannon Mattern, "Mapping's Intelligent Agents" https://placesjournal.org/article/mappings-intelligent-agents/
- Alex Galloway, "The Computational Decision" http://cultureandcommunication.org/galloway/the-computational-decision

Lab: Features and parameters; model evaluation

> Week 6 - 10/1: Project planning + regression
In the first part of this class, students will share project ideas and give each other feedback.

Readings:
- TBD: I will separately provide a list of example projects for your reference.

Lab: Regression models

> Week 7 - 10/8: NO CLASS - Midterm break

> Week 8 - 10/15: Neural networks #1
Due: Project proposal (submit by 10/13)

Readings:
- 3Blue1Brown, "Neural Networks" (YouTube playlist) https://www.youtube.com/playlist?list=PLZHQObOWTQDNU6R1_67000Dx_ZCJB-3pi

Supplemental Material:
- Daniel Shiffman, *The Nature of Code*, ch.10 https://natureofcode.com/book/chapter-10-neural-networks/
- The Coding Train, "10: Neural Networks - The Nature of Code" (YouTube playlist) https://www.youtube.com/playlist?list=PLRqwX-V7Uu6aCibgK1PTWWu9by6XFdCfh

Lab: Introduction to TensorFlow

> Week 9 - 10/22: Neural networks #2
Readings:
- Diana ben-Aaron, "Weizenbaum examines computers and society" http://tech.mit.edu/V105/N16/weisen.16n.html
- Blaise Agüera y Arcas, Margaret Mitchell and Alexander Todorov, "Physiognomy's New Clothes" https://medium.com/@blaisea/physiognomys-new-clothes-f2d4b59fdd6a

Lab: TensorFlow continued

Guest Speaker: TBD

> Week 10 - 10/29: Machine learning workflow + project workshop
Due: Lab assignment #3 (submit by 10/28)

This will be a lab-focused session. We will learn about data pipelines, re-using your models, etc. Then, we will workshop your project with the main goal of verifying that you are on track with the class timeline.

Lab: Project workshop

> Week 11 - 11/5: Pre-trained models, pre-made tools
Readings:

- Patrick Hebron, Rethinking Design Tools in the Age of Machine Learning https://medium.com/artists-and-machine-intelligence/rethinking-design-tools-in-the-age-of-machine-learning-369f3f07ab6c
- Jeremy B. Merrill, A Crash Course for Journalists In Classifying Text with Machine Learning https://qz.ai/a-crash-course-for-journalists-in-classifying-text-with-machine-learning/

Supplemental Material:

- Weka https://www.cs.waikato.ac.nz/ml/index.html
- Wekinator by Rebecca Fiebrink http://www.wekinator.org/

Lab: Introduction to Runway ML

> Week 12 - 11/12: Machine learning as a corporate infrastructure
Readings:

- Katharine Schwab, “The Dead-Serious Strategy Behind Google’s Silly AI Experiments” https://www.fastcompany.com/90152774/the-dead-serious-strategy-behind-googles-silly-ai-experiments
- Mark Bergen, “Google Wants to Train Other Companies to Use Its AI Tools” http://web.archive.org/web/20171101212341/https://www.bloomberg.com/news/articles/2017-10-19/google-wants-to-train-other-companies-to-use-its-ai-tools

Supplemental Material:

- Nick Seaver, “Knowing Algorithms” https://digitalsts.net/essays/knowing-algorithms/
- Geraldine Juárez, “A Pre-emptive History of the Google Cultural Institute” – via LMS

Lab: Runway ML continued

> Week 13 - 11/19: Unsupervised learning #1
This will be a lab-focused session that introduces unsupervised learning, which is a different ML paradigm than supervised learning (such as classification and regression).

Lab: Clustering algorithms

Guest Speaker: TBD

> Week 14 - 11/26: Unsupervised learning #2
Due: Lab assignment #4 (submit by 11/25)

This will be a lab-focused session that continues on the theme of unsupervised learning.

Lab: Dimensionality reduction; visualization

> Week 15 - 12/3: Maintenance and human labor in machine learning
Readings:

- Kate Crawford and Vladan Joler, Anatomy of an AI System https://anatomyof.ai/
- West, S.M., Whittaker, M. and Crawford, K. (2019). *Discriminating Systems: Gender, Race and Power in AI*. AI Now Institute. https://ainowinstitute.org/discriminatingsystems.pdf
- M.C. Elish and Tim Hwang, An AI Pattern Language, ch.2 (16-34) https://datasociety.net/pubs/ia/AI_Pattern_Language.pdf

Supplemental Material:

- Shannon Mattern, "Maintenance and Care" https://placesjournal.org/article/maintenance-and-care/

Lab: TBD

> Week 16 - 12/10: Presentations
Due: Final project (before class)

Textbooks, Readings and Materials:

All reading materials and course slides (if applicable) will be provided as hyperlinks or downloadable files through LMS.

Students will need a Google account for certain lab sessions. I believe the Pratt email address can serve this purpose, giving you access to Google Drive and Colab. In the latter part of the course, students will also need an account for Runway ML; details on how to sign up will be provided as needed.

Projects, Papers and Assignments:

> Readings and Discussions
Throughout the semester, we will survey diverse perspectives about machine learning as a socially situated technology. The assigned readings will be complemented by in-class discussions, typically at the beginning of the class.

Each week (except for weeks with no readings assigned; see course schedule), one or two students will act as motivators and write provocations on the readings of the week on the LMS forum. This will allow us to start the conversation in advance of class and carry it on afterwards. Please post your provocations by the end of the day Sunday before class.

A provocation will include a summary of key points in the readings, as well as questions / observations you would like to raise or make. The provocations will serve as starting points of the in-class discussion and some of them will scaffold towards the project proposal and final project.

Students who are not motivators for the week are expected to complete the readings before class, and contribute to the discussion in class and/or online by replying to the forum thread.

> Lab assignments
The latter part of each class will be a lab session related to the topic of the week. Sometimes, the lab session will be accompanied by a lecture-style session before it; in other cases, we will move into the lab session right after discussions.

At the end of some lab sessions, I will give you take-home assignments (4 total). The assignments will scaffold towards the final project. For example, you will be asked to explain some machine learning terminology or write code that does a specific task. Details on how to submit the homework will be

communicated in class. The homework assignments are due by the end of the day Monday before the next class.

I may also ask you to write down the amount of time you spent working on the assignment. This amount of time does NOT affect gradings in any way; I am asking in order to gauge whether I am giving you too much work or whether you are having trouble with some of the course material.

> Project proposal
I will ask you to choose a topic that you would like to explore in your final project, and to submit a proposal by mid-semester.

On October 1, we will have an in-class activity where you will share your idea(s) and give peer feedback. Your 800- to 1200-word proposal is due by the end of the day Sunday, October 13, and should include:

- A description of the data you intend to use
- A description of the machine learning task you intend to perform
- A tentative and brief survey of existing work on the topic
- A discussion of the significance of your topic

We will discuss the proposal in further detail in the coming weeks.

> Final project + presentation
Your final project is to run an experiment that applies a machine learning technique (such as classification, regression, clustering, etc) that we learned on a dataset of your choice. You can design the project as a complete piece on its own, or as a component of a larger project.

Projects are due before the final class in the form of a write-up detailing your work process; you also need to submit the resulting model / dataset and code used. We will dedicate our class on December 10 to presentations.

A detailed rubric for the project and presentation will be distributed separately.

Assessment and Grading

Lab assignments 20%
Participation (discussions and peer feedback) 20%
Project proposal 20%
Final project 30%
Presentation 10%

Pratt's grading scale:

Superior work:	A 4.0 (96-100)	A- 3.7 (90-95)	
Very good work:	B+ 3.3 (87-89)	B 3.0 (83-86)	B- 2.7 (80-82)
Marginally satisfactory:	C+ 2.3 (77-79)	C 2.0 (73-76)	
Failed:	F 0.0 (0-72)		

Portfolio

Work completed for this course may be included in your portfolio. For more information on each program's portfolio requirements, please visit the program's respective webpage:

- MS Library & Information Science: Portfolio - http://bit.ly/prattmslisportfolio

- MS Information Experience Design: Portfolio - http://bit.ly/prattmsixdportfolio
- MS Data Analytics and Visualization: Portfolio - http://bit.ly/prattmsdavportfolio
- MS Museums and Digital Culture: Portfolio - http://bit.ly/prattmsmdcportfolio

You are encouraged to meet with your advisor about including projects in your portfolio.

Policies

This Course's Attendance Policy

Students are allowed two individual absences for any reason; no notice or documentation is required. If you miss a session, be sure that you complete the readings, consult your classmates or the professor about the class, and (as always) demonstrate your knowledge of previous topics in later sessions. Upon a third or fourth absence, your grade will be lowered by one mark for each absence. Five absences will result in failure of the course (F). Also, please note that a grade of C- is considered failing at the graduate level.

Students who encounter long-term health issues should provide documentation to the Office of Health and Counseling and discuss their options with the professor.

Absences	**Penalty**	**Examples**
0-2	None	N/A
3-4	One mark per absence	A → A- (3rd absence) → B+ (4th absence) A- → B+ (3rd absence) → B (4th absence) B+ → B (3rd absence) → B- (4th absence)
5+	Failure of the course	N/A

For more information on Pratt Institute's Attendance Policy, please visit http://bit.ly/prattattendance.

Academic Integrity Code

Academic integrity at Pratt means using your own and original ideas in creating academic work. It also means that if you use the ideas or influence of others in your work, you must acknowledge them. For more information on Pratt's Academic Integrity Standards, please visit http://bit.ly/prattacademicintegrity.

Students with Disabilities and Accessibility

Pratt Institute is committed to the full inclusion of all students. If you are a student with a disability and require accommodations, please contact the Learning/Access Center (L/AC) at LAC@pratt.edu to schedule an appointment to discuss these accommodations. Students with disabilities who have already registered with the L/AC are encouraged to speak to the professor about accommodations they may need to produce an accessible learning environment.

Requests for accommodation should be made as far in advance as reasonably possible to allow sufficient time to make any necessary modifications to ensure the relevant classes, programs, or activities are readily accessible. The Learning/Access Center is available to Pratt students, confidentially, with additional resources and information to facilitate full access to all campus programs and activities and provide support related to any other disability-related matters.

For more information, please visit http://www.pratt.edu/accessibility/.

Revisions to the Syllabus

All changes will be communicated in class and the updated syllabus will be uploaded on LMS. This version is current as of August 26, 2019.