CSc 44700 – Introduction to Machine Learning
Fall 2021
Erik K. Grimmelmann, Ph.D.

Syllabus v1.0

Overview

This course will provide a theoretical and hands-on introduction to the basics of machine learning and its application to various real-world problems.

Prerequisites

You are expected to have a basic knowledge of computer science including object-oriented programming, algorithms, data structures, and data analysis and a working knowledge of Python. In addition, you are expected to have a rudimentary understanding of probability, statistics, and basic linear algebra.

The formal prerequisites are

- CSc 21200 Data Structures
- CSc 21700 Probability and Statistics for Computer Science
- CSc 22000 Algorithms
- CSc 22100 Software Design Laboratory
- CSc 30100 Scientific Programming
- Math 34600 Elements of Linear Algebra

If you have not taken (and done well in) all of these courses (or their equivalents), please check with me prior to continuing in the course.

Class Meetings via Zoom

We will meet Mondays and Wednesdays

- Section E 2:00 pm to 3:15 pm

Please note that, per the CCNY calendar, no classes for this course are scheduled on the following Mondays or Wednesdays:

Monday, September 6
Wednesday, September 8
Wednesday, September 15
Monday, October 11

Attendance in class is required.

To receive credit for attending, you must post "present" in the Zoom chat.

If you will miss a class for a good reason (e.g., illness or a job interview), please email me; I may give you some partial credit for attendance that day provided that you watch the recording of the Zoom session. Your attendance and class participation will be an important component of your final grade.

Please **arrive promptly**. We will be starting class at the appointed time. Arriving late is better than not arriving at all. Your arriving on time will be a component of your final grade.

Please **pay attention** in class. **No texting or doing anything else on-line during class**

Zoom

We'll be using Zoom for our distance learning this semester. The Zoom links for the classroom sessions and my office hours are posted in Blackboard.

Here are some tips for using Zoom:

- Use an up-to-date version of Zoom.
- Make sure that the name that appears indicates is recognizable as your name. If you need to change the name that appears, right-click on your name on the Zoom screen, choose Rename, enter your new screen name, and hit enter. If I can't connect your screen name with you, you won't get credit for attending the class meeting.
- Make sure to sign in by typing "present" in the Zoom chat. If you don't do this during the class session, you won't get credit for attending.
- Mute your audio except when you want to say something to the class.
- Speak up when you do have something to say. It will be a lot more interesting for all of us if I'm not the only one speaking.
- If I'm speaking and you can't hear me, please let me know via chat and/or speaking up; I may have unintentionally muted my microphone.

- Remind me (via chat and/or speaking up) if we've gone over the allotted time. I've been known to get carried away with the material and lose track of the time.
- Leave the session at the end of the class; otherwise, I may think that, although you've been in the session, you've left to do something else and haven't been paying attention.

I will record the class session and post them on Blackboard in the "Zoom Recordings" tab within a day or two of each class session. Due to storage limitations on Blackboard, recordings will be online for only two weeks or so.

Textbook

The required textbooks are

- Raschka, Sebastian, and Vahid Mirjalili. *Python machine learning: machine learning and deep learning with Python, scikit-Learn, and TensorFlow*. Packt, 2017. Note that this is the second edition of this book; the first edition had only Raschka as the author and a slightly different title.

- Chollet, Francois, *Deep learning with python*, Shelter Island, NY: Manning, 2018.

In general, the lectures will not be drawn from the textbook but rather from a number of other works in the field. I'll be letting you know which sources I've used for the topics we cover.

A bibliography with suggested other references will be posted on the Blackboard site.

Guest Lectures

Depending on their availability, I will try to bring some of my friends and mentors in to speak to the class. I'll let you know when guest lectures are scheduled.

Programming Exercises

All programming in the course will be in Python and its relevant add-ins and libraries (such as NumPy, SciPy, MatPlotLib, Pandas, and Scikit-learn) and TensorFlow. You are free to work in any environment that supports Python (e.g., Windows, Max, Unix, Linux). We'll be using Jupyter notebooks throughout the course. Programming exercises will be an important component of your final grade

Final Project

During the final third of the semester, you'll work on an individual project.

The project will include

- A Zoom presentation with charts (in PowerPoint, Google Slides, etc.) The presentation will be six to seven minutes long.
- A written report (in PDF format). The report will be 10 to 15 pages in length (double spaced). If you have lots of tables and/or charts, you can include them in an appendix.

I'll be providing more detail about the project later. The final project will be an important component of your final grade.

Blackboard

We will be using Blackboard as our online environment. Once you're enrolled in the course and the course has started, you should have access to the Blackboard course site. We will use the course site for

- This syllabus
- Links to reference materials
- Announcements
- Posting and submission of assignments
- Classroom presentations (typically within a week of the class session).
- Zoom recordings of the Zoom classroom sessions
- Datasets
- Sample code
- Bug reports
- Assignment grades

Final course grades will not be posted on Blackboard, but rather on CUNYfirst.

Course Policies

Collaboration: Except where I tell you otherwise, you are free to collaborate freely with each other and to consult any sources you wish to in your work for this class. If you collaborate on an assignment, you must have contributed substantially to anything you submit for the

assignment; just using a (current or past) classmate's code without having contributed substantially to it is not collaboration; it's cheating.

If you collaborate on an assignment, you most indicate in your submission with whom you collaborated.

Professionalism: I expect you to act professionally and respectfully to your classmates to me (and to our occasional guests, if we have any) at all times. Harassment will not be tolerated.

If for any reason your preferred name is not the one that appears on the course roster, please let me know how you would like to be addressed.

Grades

Your grades will be based on the following factors:

- Classroom activities 25%
  - Attendance
  - Punctuality
  - Participation
  - Bug bounties (extra credit)
- Assignments including programming exercises 40%
- Individual final project 35%

Integrity

Just to refresh your memory, here's the City College statement on academic integrity:

> Academic integrity is an essential part of the pursuit of truth, and of your education. We are all are all responsible for maintaining academic integrity at City College – it is the rock on which the value of your degree is built.
>
> If you cheat on a test or plagiarize by using someone else's work or ideas, you defeat the purpose of your education. In addition, academic dishonesty is prohibited in the City University of New York, and is punishable by failing grades, suspension and expulsion.

Here's a link to a list of City College and CUNY policies (and links to them), https://www.ccny.cuny.edu/about/policies

If you use code from any source other than your own imagination for any coding assignment, be sure to list the source(s). If you collaborate with someone else on a coding assignment, you must indicate with whom you collaborated in your submission.

Your feedback

I welcome your feedback at all points in the course. If something is unclear, please speak up. If you find an error in my lectures, code examples, assignments, or in anything else, please point it out and submit a bug report.

My Contact Information

The best (and fastest) way to reach me is via email at egrimmelmann@ccny.cuny.edu

Office Hours

My office hours will Mondays and Wednesday, from 12:30 pm to 1:30 pm on days that we have class. There will be a Zoom link for my office hours These office hours will be shared with the Scientific Programming and Senior Project courses that I'm also teaching this semester.

CCNY Resources

Here are links to some of the resources that are available to you.

- AccessAbility https://www.ccny.cuny.edu/accessability
- Health and Wellness https://www.ccny.cuny.edu/health-wellness
- Student Affairs https://www.ccny.cuny.edu/studentaffairs
- Writing Center https://www.ccny.cuny.edu/writing

Course Schedule

The schedule below is almost certainly more precise than it will be accurate. We'll likely end up going faster on some of the topics and slower on others, so we could end up being ahead or behind of this schedule at varying points in the course. Please read the material (i.e., the chapter in the textbook or posting on Blackboard) prior to our covering the material in class.

Meeting 1	Wed	Aug 25	Course introduction Context for the growth in machine learning
Meeting 2	Mon	Aug 30	Introduction to machine learning
Meeting 3	Wed	Sep 1	The Perceptron

No Class	Mon	Sep 6	**No CCNY Classes**	
	Wed	Sep 8		
Meeting 4	Mon	Sep 13	Adaline Model, Multiclass Classification	
No Class	Wed	Sep 15	**No CCNY Classes**	
Meeting 5	Mon	Sep 20	Logistic Regression	
Meeting 6	Wed	Sep 22	Support vector machines, Lagrange multipliers	
Meeting 7	Mon	Sep 27		
Meeting 8	Wed	Sep 29	Decision trees	
Meeting 9	Mon	Oct 4		
Meeting 10	Wed	Oct 6	Bayesian learning	
No Class	Mon	Oct 11	**No CCNY Classes**	
Meeting 11	Wed	Oct 13	Confusion matrices, ROC curves	
Meeting 12	Mon	Oct 18	Strong & weak learner, boosting, random forests	
Meeting 13	Wed	Oct 20	Regression	
Meeting 14	Mon	Oct 25		
Meeting 15	Wed	Oct 27	KNNs, principal component analysis	
Meeting 16	Mon	Nov 1	Neural networks	
Meeting 17	Wed	Nov 3		
Meeting 18	Mon	Nov 8	Convolutional neural networks, computational graphs	
Meeting 19	Wed	Nov 10	Recurrent and recursive neural networks	
Meeting 20	Mon	Nov 15	Reinforcement learning	
Meeting 21	Wed	Nov 17	Ethics and fairness in machine learning	
Meeting 22	Mon	Nov 22		
Meeting 23	Wed	Nov 24	Generative models	
Meeting 24	Mon	Nov 29	Guest Lecture	
Meeting 25	Wed	Dec 1	Project presentations	
Meeting 26	Mon	Dec 6		
Meeting 27	Wed	Dec 8		
Meeting 28	Mon	Dec 13		
No Class	Wed	Dec 15		**CCNY Exam Period**
	Mon	Dec 20	Project presentation charts due	
	Tue	Dec 21	Written projects due	