

Machine Learning for Finance Syllabus (Spring 2023)

(Subject to change)

Course Overview

The application of Machine Learning (ML) algorithms in the Financial industry is now commonplace, but still nascent in its potential. This course provides an overview of ML applications for finance use cases including trading, investment management, and consumer banking. Students will learn how to work with financial data and how to apply ML algorithms using the data. In addition to providing an overview of the most commonly used ML models, we will detail the regression, KNN, NLP, and time series deep learning ML models using desktop and cloud technologies. The course is taught in Python using Numpy, Pandas, scikit-learn and other libraries. Basic programming knowledge in any language is required.

Requirements

- Basic programming skills in any language, familiarity with Python is preferred.
- No ML knowledge required
- Finance and math knowledge, primarily basic calculus and linear algebra.

Grading

- 5 homeworks, 40% of grade
 - Grade best 4 out of 5
 - Mostly python code application of concepts
- 2 projects, 20% of grade
 - 1 individual project
 - 1 team project, random team memberships or organize your own
- 2 exams, 40% of grade
 - Closed book, given online using Proctorio
 - Mostly concept questions, but some questions will require python knowledge/application

Learning Objectives

1. Understand how to access, cleanse, and analyze data using application programming interfaces (API) in Python.
2. Get an overview of the most common machine learning algorithms and applications.
3. Deep dive machine learning algorithms for financial use cases, including trading, investment management, and consumer banking.
4. Learn how to use Python libraries like NumPy, Pandas, scikit learn, and Beautiful Soup.
5. Learn basics of ML cloud technologies

Readings

All optional:

1. [Machine Learning in Finance, Dixon, Halperin, Bilokon](#) (MLF)
2. [Python for Finance: Mastering Data-Driven Finance 2nd Edition](#) (PF)

3. [The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition by Hastie, Tibshirani, Friedman \(ESL\)](#)
4. [Python for Data Analysis by Wes McKinney \(PDA\)](#)
5. [Machine Learning by Tom M. Mitchell 1st Edition \(ML\)](#)
6. [An Elementary Introduction to Mathematical Finance 3rd Edition by Sheldon M. Ross \(EIMF\)](#)

Modules

1. Working with Data
2. Platform APIs
3. ML Algorithms
4. ML Applications in Finance
5. Cloud Technologies for Finance

Course Schedule

Week	Module	Topic	Readings	Assignments Due
1 1/21	1	1. Course Overview, Motivation	PF ch. 1,3	
2 1/28	1	2. ML Intro 3. Framing the ML Problem 4. Data Collection 5. APIs – REST, SQL	PF ch. 5,6,7	HW0 – not graded
3 2/4	2	6. Generalization 7. Data Split		
4 2/11	3	8. Loss, Regularization		HW1
5 2/18	3	9. Classification	ESL ch. 3,4	
6 2/25	3	10. Clustering	ESL ch. 13	HW2
7 3/4	Exam	Midterm		
8 3/11	4	11. <SPRING BREAK>		
3/18		NLP	Link	HW3
9 3/25	4	12. Neural Nets		
10 4/1	4	13. Neural Nets		HW4
11 4/8	4	14. Time Series		Project 1
12 4/15	5	15. ML on the Cloud	GCP AWS	
13 4/22	5	16. Reinforcement Learning		HW5 (optional)
14 4/29		17. Project Presentations		Project 2
15 5/6	Exam	Final		

Class and University Policies

Academic Integrity

Columbia's intellectual community relies on academic integrity and responsibility as the cornerstone of its work. Graduate students are expected to exhibit the highest level of personal and academic honesty as they engage in scholarly discourse and research. In practical terms, you must be responsible for the full and accurate attribution of the ideas of others in all of your research papers and projects; you must be honest when taking your examinations; you must always submit your own work and not that of another student, scholar, or internet source. Graduate students are responsible for knowing and correctly utilizing referencing and bibliographical guidelines. When in doubt, consult your professor. Citation and plagiarism-prevention resources can be found at the GSAS page on [Academic Integrity and Responsible Conduct of Research](#).

Failure to observe these rules of conduct will have serious academic consequences, up to and including dismissal from the university. If a faculty member suspects a breach of academic honesty, appropriate investigative and disciplinary action will be taken following [Dean's Discipline procedures](#).

Disabilities Accommodations

If you have been certified by Disability Services (DS) to receive accommodations, please either bring your accommodation letter from DS to your professor's office hours to confirm your accommodation needs, or ask your liaison in GSAS to consult with your professor. If you believe that you may have a disability that requires accommodation, please contact [Disability Services](#) at 212-854-2388 or disability@columbia.edu.

Important: To request and receive an accommodation you must be certified by DS. Additional information regarding Disability Services can be found [here](#).