Course: AI in Asset Pricing

QFGB-890P (49754)

Spring 2025, January 14 ~ May 6

Instructor: Dr. Ren-Raw Chen
Room: 140W-213
Class: 3:30 pm ~ 5:30 pm, Tuesdays
Office: 609 Martino Hall
Website: https://faculty.fordham.edu/rchen/teach_ai.html
Office phone: 212-636-6471
E-mail: rchen@fordham.edu
Office hours: 11:30 am ~ 12:30 pm and 5:30 pm ~ 6:30 pm, Wednesdays
TA: Karandeep Sonewane <kcs@fordham.edu>

Sessions: 1/14, 1/28, 2/11, 3/4, 3/25, 4/8, 4/22, 5/6

A Few Words

The objective of this course is to use AI/ML to evaluate complex financial (often derivatives) contracts. To do that, we not only need to be able use on-the-shelf libraries, but also need to understand the underlying math and algorithms. In other words, we will "open the black box" of every method we use in this class. Derivations as well as good skills in Python or R are required for this course.

Text

Hull Machine Learning in Business: An Introduction to the World of Data Science Lecture Notes

Grading Policy

This class is project-driven. Each project will be weighed equally, totaling to 70% of the grade (i.e. each project is 10%). The final exam which is 30% is all math.

Objective and Grading Policy

This course is predominantly a math course with a substantial emphasis on lab work (via homework assignments). The math part is to help you gain a competitive edge over those who only know how to program but don't know why. In other words, I will try my best to bridge finance and computer science (this is similar to my finance theory course where I bridge finance and probability theory). It is quite pleased to know that many AI models are not necessarily a black box.

Although this is my second time already teaching this course, there is still a great amount of uncertainty in both the materials and the style of this course. I change quite a lot from last year so it is still quite an experiment. I rely on your feedback to improve this course.

There will be 7 lectures (1/21, 2/4, 2/25, 3/11, 4/1, 4/15, and 4/29) and a mandatory final exam (5/6). The final exam will account for 40% of the grade and 6 homework assignments 60%.

Tentative Schedule

Here are the topics I intend to cover:

- graph theory (2 ~ 3 lectures)
- Bayesian learning and multi-variate normality (1 ~ 2 lectures)
- swarm intelligence and PSO (½ ~ 1 lecture)
- reinforcement learning (1 lecture)
- regularization (1 lecture)
- NLP (1 lecture)
- NN (1 lecture)
- GPU computing (½ ~ 1 lecture)
- other topics – classification, decision tree, filtering, etc.

Assignments and Descriptions (will be updated after each class)

I. Bayesian learning (Chapter 1 – 1.6 & Chapter 4 – 4.4)

1. Naive Bayes
2. Black-Litterman model

Reading (the underlined are must):

- "On the Black–Litterman Model: Learning to Do Better" by Ren-Raw Chen, Shih-Kuo Yeh, and Xiaohu Zhang, JFDS, 2022. https://faculty.fordham.edu/rchen/JFDS-Chen.pdf
- "Black Litterman Asset Allocation Framework – 01" by William Huang 2019. https://kaihuahuang.github.io/2019/09/09/Basic-Black-Litterman-Asset-Allocation/?from=timeline&isappinstalled=0
- "A Step-by-Step Guide to the Black-Litterman Model" by Thomas M. Idzorek 2004 https://people.duke.edu/~charvey/Teaching/BA453_2006/Idzorek_onBL.pdf
- "Asset Allocation: Combining Investor Views with market Equilibrium" by Fischer Black and Robert Littermen, JFI, 1991.
- https://machinelearningmastery.com/naive-bayes-classifier-scratch-python/ by Jason Brownlee

Homework: data https://faculty.fordham.edu/rchen/pima-indians-diabetes.data.csv

- repeat what Jason Brownlee did. due on 2/11, HARDCOPY only

Project: Black-Litterman – due on 1/28, HARDCOPY submission only

- Collect 1 year of daily prices of 4 financial assets (stocks, ETFs, crypto, anything) and the S&P 500 index.
- Specify any P with at least two views (must have at least 1 absolute view and at least 1 relative view)
- ~~Solve the Markowitz weights and compute its Sharpe ratio and solve for the weights of the BLM by maximizing the Sharpe ratio.~~
- Outputs in the report: (1) two sets of weights and (2) two Sharpe ratios, and (3) include the above 2 in a summary and also include a mention of data and methodology (which is BLM) – basically write an investment report.

- Bonus 1: Follow my paper and run a time series (rolling window) of the above. Drawdowns are nice too.
- Bonus 2: Follow the paper to "learn" the view.

II. Undirected Graph

1. Topology
2. Gaussian graphic models (Baysian network)

Reading (all of them are must):

- Leonard Euler's 7 (Königsberg, Kaliningrad today) bridge theorem https://www.youtube.com/watch?v=nZwSo4vfw6c&t=35s
- Knowledge graph of systemic risk by Chen and Zhang (https://faculty.fordham.edu/rchen/JFS.pdf)
- Network Analysis and Visualization with R and igraph
- https://faculty.fordham.edu/rchen/scikit-learn1.1.2.documentation.pdf

Homework

Project: Knowledge graph – due on 3/4, HARDCOPY submission only

- For the project you need volatility which you can calculate easily using historical return data. Let's use daily returns and let volatility be past one year's. Once you have times series of volatility of many stocks (minimally 20), perform graphic LASSO and draw a knowledge graph. Read my paper above for the details.

III. Directed Graph

1. Causality
2. Estimation (pc-algorithm)

Reading

- https://www.bradyneal.com/Introduction_to_Causal_Inference-Dec17_2020-Neal.pdf by Brady Neal
- Term structural of interest rates with graphic model by CHHL
- DoWhy
- IDA

Homework

Project: DAG for stocks – due on 3/25, usual rules apply.

- Collect data (returns), either individual stocks, indices, interest rates, or anything you like.
- Follow CHHL (https://faculty.fordham.edu/rchen/DAG_CHHL.pdf) and run pc-algorithm and obtain a CPDAG (you can just read the introduction, Graphic Model (which is a review of the last class), and DAG Estimation (up till equation (1.4)).
- Bonus – you can continue on to estimate edges (as in the paper) using regression (as you can see that now muti-variate normality is assumed, which links back to Chapter 2 – told you! multi-variate normal is everywhere!)

IV. NLP (Chapter 9)

1. LDA (Latent Dirichlet allocation)
   - data: https://github.com/AvinashTiwari/Artifical-intelligence/blob/master/7_nlp_python/19_Latent_Dir/npr.csv
3. Transformer

Reading

Homework due on 4/8

- Run Alina's example (https://faculty.fordham.edu/rchen/npr.csv)
- Hull homework #9.13 on page 211

Project: Topic modeling with stocks – due on 4/8

- Collect stock returns data and run LDA (the data should include at least 20 stocks and 40 periods (weeks or months))
- Compare it with PCA (optional)

V. Reinforcement Learning (Chapter 8)

1. Q learning
2. Markov decision process
3. Longstaff-Schwartz model

Reading

- https://people.math.ethz.ch/~hjfurrer/teaching/LongstaffSchwartzAmericanOptionsLeastSquareMonteCarlo.pdf

Homework due on 4/22

- Replicate the example of Longstaff-Schwartz
- (bonus) Write an **executable** Python or R program for the "maze" problem. Show the optimal path. Submit the program to Sonewane.

Project due on 4/22

- Implement the Longstaff-Schwartz with a real option example (e.g. 10,000 paths, 100 periods, Black-Scholes stochastic process with r=3% and σ=0.4, etc.) I posted a paper of mine using GPU for your reference. There you will find discussions of more efficient basis functions. You can also see how to use GPU (which we may not have time to discuss). Hope it interests you: https://faculty.fordham.edu/rchen/gpu.pdf
- Implement RL with the same sample path (and assume the same quadratic basis function)

VI. Swarm Intelligence

1. Boids
2. PSO

Reading

- "An Artificial Intelligence Approach to the Valuation of American-style Derivatives: A Use of Particle Swarm Optimization," with William Huang, Jeffrey Huang, and Robert Yu, Journal of Risk and Financial Management, Vol. 14, No. 2, 2021.
- "Index Tracking: A Stock Selection Model using Particle Swarm Optimization," Journal of Investing, 32(2), 2023.
- "Classifying Factor Velocity with Swarm Intelligence: Market Pricing of Fast- and Slow-Moving Factors" with Yi Tang.

Homework – submit program for testing

- "Artificial Intelligence Approach to the Valuation of American-style Derivatives: A Use of PSO" by Chen et. al. JRFM, 2021 (https://faculty.fordham.edu/rchen/jrfm-14-00057.pdf) due on **4/29 (Note that this is electronic submission. Please email an executable program to Mr. Sonewane.)**

- (bonus) Traveling salesman problem

Project

- Use the Fama-French 3-factor model to replicate the "Classifying Factor Velocity with Swarm Intelligence: Market Pricing of Fast- and Slow-Moving Factors" by Chen and Tang (https://faculty.fordham.edu/rchen/paper_algorithms.pdf). The data can be found on French's website: https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html due on **4/29 (paper submission)**

VII. Clustering (Chapters 3 and 4)

1. Probit/Logit, Discriminant
2. k-means, etc.

Reading

Homework

-

Project

VIII. GPU/Quantum Computing