Data Science in Quantitative Finance and Risk Management
2022 Summer Immersion Program
The University of Chicago

Course Overview

Current Grade/ Educational Level: 10th Grade,11th Grade, 12th Grade

Program Type: Summer Online

Start Date: July 5, 2022 (Tuesday)

End Date: August 5, 2022 (Friday)

Summer Session Link:
https://summer.uchicago.edu/course/data-science-quantitative-finance-and-risk-management

Class Details

Instructor: Yin Kwong (John) Lee - lykjohn@uchicago.edu

Course Assistants:

- Shunqi (Johnson) Zhang - shunqi@uchicago.edu
- Seung Chul (Eric) Lee - ericsclee@uchicago.edu

Office Hours: Please email the course instructor/assistant to set an appointment.

Academic Interest: Mathematics, Statistics, Finance, Economics, Business, Computer Science

Class Specifics

Course Code: STAT 13820 94

Class Days & Time: Mon, Tue, Wed, Thu, Fri 6:00 PM-8:00 PM CDT

Session: Session II

Course Length: 5 Weeks

Course Description

This course aims to introduce leading statistical models and methods for financial data research, in which learners will utilize techniques in statistical modeling, machine learning, and risk management to build insightful financial strategies that best suit their personalities. In the beginning, students will gain hands-on exposure to exploring real-time market data. Then, they will practice risk analysis with proprietary metrics such as Value-at-Risk(VaR) used to evaluate portfolio returns/losses. After that, students will learn about the theoretical and applied foundations of regression and classification designs for predicting market patterns. Finally, they will be introduced to statistical learning under a Bayesian framework. These techniques can be applied to the U.S. and foreign asset classes, including equities, commodities, and cryptocurrencies. All implementations are done using Python. Students will form research teams to play a stock market game using the skills they learned throughout the course

with the objective of building their desired portfolios. Teams will get a chance to present and share their investment journey with the class towards the end of the course.

Prerequisites

Experience with at least one programming language (e.g., Python, R, C++) is required. Prior knowledge of statistics and probabilities is helpful but not required. Knowledge of financial markets is not required.

References

- Stefan Jansen, *Machine Learning for Algorithmic Trading: Predictive models to extract signals from market and alternative data for systematic trading strategies with Python, 2nd Edition (July 31, 2020)*
- Abdullah Karasan, *Machine Learning for Financial Risk Management with Python: Algorithms for Modeling Risk 1st Edition (January 18, 2022)*

Grading

- Lab assignments: 60%
- Team presentation: 40%
- Discussion/ Participation bonus: up to 5%

Letter Grade Translator

(all range percentages are inclusive)

Grade	**Range**
A+	94%-100%
A	90%-93%
A-	87%-89%
B+	84%-86%
B	80%-83%
B-	77%-79%
C+	74%-76%
C	70%-73%
C-	67%-69%
D+	64%-66%
D	61%-63%
F	<60%

Lab: 60%

- There will be weekly lab assignments. Late lab is accepted for up to 4 days after the due date, with a (# of days)*25% reduction for each late day. Typically, a lab will consist of data analysis topics covered in class that week, to be done in with Python in Jupyter notebook.
- You are encouraged to discuss labs with each other, but ultimately all questions must be solved and written up independently. Identical lab responses will receive no credit and lead to notable consequences. Any form of plagiarism (including from open online sources) in the labs will be taken very seriously and can result in failing the course.
- In case of medical reasons for a late lab, a doctor's note is needed at least 48 hours before the lab is due.
- Please direct any questions about the lab to John, Johnson, Eric, or ask them on the discussion board (you might get faster answers there).

Team Presentation: 40%

- Each team will participate in a stock market game, a collaborative space for them to journal the evolution of their investment portfolios as different analyses are applied throughout the course.
- In real life, quants, consultants, and asset managers most often perform similar tasks on a day-to-day basis, presenting their results to portfolio owners. Imagine you are one of them.
- As a team, you will select a portfolio of your choice at the beginning of week 1, then after applying the methods learned from week 1, journal any changes to the original portfolio. This will be your starting portfolio in week 2, then after applying the methods learned from week 2, journal any changes to the portfolio. Repeat the process until your team presentation.
- As a guide, the presentation should not be longer than 5 minutes for each team, and should address the following:
    1) What does the portfolio look like before and after the weekly analysis?
    2) What types of analysis are performed?
    3) How do the analysis results affect your investment decisions?
    4) How do the return and risk of your portfolio look like along the way?
    5) How else can the team help to improve the client's portfolio return?
- Each team must schedule a weekly 15-minute time slot to check in with an instructor
- After each presentation, there will be 2 mins open to the class for Q&A.
- All team members are supposed to work on the presentation together and each member must speak to their parts in the presentation.
- A single grade is given to all members of a group. Late presentation, just like late labs, is accepted for up to 2 days, with a (# of days)*50% reduction per late day.
- In case of medical reasons for a presentation, a doctor's note is needed at least 48 hours before the presentation is due.
- Please direct any questions about the team presentation to John, Johnson, Eric or ask them on the discussion board (you might get faster answers there).

Discussion/ Participation Bonus: up to 5%

- There will be weekly discussion boards, where individual students can post, comment, or ask questions to show efforts of participation.
- Discussion topics can freely range from recommended readings to course contents to market news and climates, and even to career aspirations.

Computing Technology & Tools

- Computer with WIndows 10+, macOS. Linux
- Stable internet because we will be pulling live data from sources like yahoo finance
- We will be using packages, particularly scikit-learn and others, in Python to build machine learning models from financial data. Python can be downloaded for free from https://www.python.org/. If you prefer to use a different programming tool, you are welcome to use it, but sample code will only be provided for Python
- Here is the guide to access Python in Jupyter Notebook.

Course Outline

Date	Lecture Materials	Assignments

July 5th-July 8th (Week 1)	• Students form teams for the stock market game. Every team can edit their portfolios after logging into the Stock Market Game platform: https://www.stockmarketgame.org/login.html
	• Overview of the quantitative industry (e.g., algorithmic vs discretionary technologies)
	• Set up Python and Jupyter Notebook on the computer, install relevant packages
	• Review on major Python packages such as numpy, pandas, matplotlib, scikit-learn, etc.
	• Connect to the yahoo finance data source and explore the data structures (e.g., high, low, close, volume)
	• Exploratory Data Analysis (e.g., correlation heatmap,
		• Week 1 Lab: Perform Exploratory Data Analysis (EDA) for an asset of your choice- due by 11:59pm CDT on July 10th
		• 1st team presentation check-in - meet by 11:59pm CDT on July 11th
		• Week 1 Reading:
		  ○ A Walk-through of Principal Component Analysis
		  ○ Cluster Analysis – Machine Learning for Pairs Trading

	label balancing, summary statistics, histogram, normality measures)	
July 11th-July 15th (Week 2)	• Introduction to probability distributions (e.g., normal, student-t, double-exponential) • Tail probability calculation used to derive portfolios risk metrics (e.g., VaR, Expected Shortfall)	• Week 2 Lab: Practice portfolio hedging and risk management - due by 11:59pm CDT on July 17th • 2nd team presentation check-in - meet by 11:59pm CDT on July 18th • Week 2 Reading: Recipe for Disaster: The Formula That Killed Wall Street
July 18th-July 22th (Week 3)	• Introduction to topics in regression (e.g., simple/multiple linear, ridge, lasso, and logistic regressions) • Goodness of fit measures (e.g., R-squared, residual measures, diagnostic plots) used to evaluate model performances	• Week 3 Lab: Implement machine learning methods and models on regression problems - due by 11:59pm CDT on July 24th • 3rd team presentation check-in - meet by 11:59pm CDT on July 25th • Week 3 Reading: Cross Validation — Why & How
July 25th-July 29th (Week 4)	• Introduction to machine learning topics in classification (e.g., nearest neighbor classifiers, support vector machines, decision trees) • Goodness of fit measures (e.g.,confusion matrix, mean square error, AUC-ROC curve) used to evaluate model performances	• Week 4 Lab: Implement machine learning methods and models on classification problems - due by 11:59pm CDT on July 31st • 4th team presentation check-in - meet by 11:59pm CDT on August 1st • Week 4 Reading: Forecasting Time Series Data - Stock Price Analysis
August 1st-August 5th (Week 5)	• High-level introduction to Bayesian machine learning (e.g., prior, likelihood, posterior constructions) • Markov Chain Monte Carlo Sampling Methods (e., No-U-Turn Sampler) with linear models (e.g., simple,	• Team presentations - due during class time 6:00 pm - 8:00 pm CDT on August 5th • Week 5 Lab (optional): Bayesian machine learning from a high level - due by 11:59pm CDT on August 7th • Week 5 Reading: The Science of

	rolling, and logistic regressions)	Portfolio Optimization

Learning Outcomes

Through the course content, students will get familiar with the hands-on implementation of machine learning methods. Most importantly, these methods, along with traditional risk management approaches can help students confidently deploy statistical tools to understand the ever-evolving financial market. Through the team presentation, students will experience how professionals in quantitative trading, hedge funds, and risk analytics collaborate to pitch asset strategies to their clients. Through participating in the stock market game, students will develop proper instincts for real-life research and execution.

Disclaimer

Although we believe that the activities in this curriculum are beneficial to helping any learner understand the financial market from a statistical perspective, we can give no assurance that practicing such activities in the real market will prove favorable profits. Various factors could cause actual performance to differ from what was expected from the strategies taught. Factors regarding the economy, securities market, or other specialized areas, like all predictors of future events, are not guaranteed to be accurate and may lead to circumstantial losses.

Academic Conduct

When participating in a group discussion or on a discussion board, please do the following:

- Criticize ideas, not individuals.
- Commit to learning, not debating. Comment in order to share information, not to persuade.
- Avoid blame, speculation, and inflammatory language.
- Allow everyone the chance to participate.
- Avoid assumptions about any member of the class or generalizations about social groups. Do not ask individuals to speak for their (perceived) social group.

Accommodations for Students with Disabilities

If you think you need accommodation for a disability, please let me know at your earliest convenience. Some aspects of this course, the assignments, team activities, and the way the course is usually taught may be modified to facilitate your participation and progress.