Accounting Data-Driven Investment Analysis ACT6211

Instructor: Xiaoxu Li

Date: 2019/09/14

Venue: TA101

Course overview

- Modern investment analysis is increasingly quantitative in nature and investment strategies are more and more determined by accounting data-driven models. This course aims to provide student with the quantitative foundations involved in investment decisions backed with data analysis and statistical inferences. This course introduces machine learning methods involved practical applications in financial investment and financial decision making. Student will learn how to build, test, and implement the types of models in use today that are essential for key investment decisions by asset managers nowadays.

Requires

- Prerequisites: ACT 5006
- Textbook: Advances in Financial Machine Learning, Marcos Lopez de Prado
- Maybe add tools or applications that students need to use and install before the class.

Author Profile

Marcos López de Prado is the CIO of True Positive Technologies (TPT), and Professor of Practice at Cornell University's School of Engineering. He has over 20 years of experience developing investment strategies with the help of machine learning algorithms and supercomputers. He launched TPT after he sold some of his patents to AQR Capital Management, where he was a principal and AQR's first head of machine learning. Marcos also founded and led Guggenheim Partners' Quantitative Investment Strategies business, where he managed up to $13 billion in assets, and delivered an audited risk-adjusted return (information ratio) of 2.3.

Concurrently with the management of investments, between 2011 and 2018 Marcos was a research fellow at Lawrence Berkeley National Laboratory (U.S. Department of Energy, Office of Science). He has published dozens of scientific articles on machine learning and supercomputing in the leading academic journals, is a founding co-editor of The Journal of Financial Data Science, and SSRN ranks him as the most-read author in economics. Among several monographs, Marcos is the author of the graduate textbook Advances in Financial Machine Learning (Wiley, 2018).

Marcos earned a PhD in financial economics (2003), a second PhD in mathematical finance (2011) from Universidad Complutense de Madrid, and is a recipient of Spain's National Award for Academic Excellence (1999). He completed his post-doctoral research at Harvard University and Cornell University, where he is a faculty member. In 2019, Marcos received the 'Quant of the Year Award' from The Journal of Portfolio Management.

Learning outcomes

- Upon completing this course, students will be able to:

① Understanding financial data structure, such as data source representation, standard bars and sampling features.

② Acquire scientific steps in building a trading strategy, and understand the pitfalls in a strategy research and backtesting.

③ Apply machine learning algorithms / models in investment.

Course syllabus

Class	Topic	Chapter
1	Financial Data Structures	1
2	Labeling	1
3	Sample Weights	1
4	Fractional Differentiated Features	1
5	Ensemble Methods	2
6	Cross-Validation in Finance	2
7	Feature Importance	2
8	Bet Sizing	3
9	Backtesting through Cross-Validation	3
10	Backtest Statistics	3
11	Understand Strategy Risk	3
12	Machine Learning Asset Allocation	3
13	Structural Breaks	4
14	Microstructural Features	4

Assessment scheme

Component	% Weight
Homework Assignments	40%
Final Exam	60%
Total	100%

Information on instructor and tutor

Instructor	Li, Xiaoxu
Office	
Email	lixiaoxu06@gmail.com
Office hour	

Teaching Assistant	Wang, Yue (Ann)
Office	TA 517-521
Email	wangyue@cuhk.edu.cn
Office hour	Tuesday 10am-2pm

Pop quiz !!!

① Have you used the Python before?

② What is your understanding of machine learning?

③ Any difference between machine learning and financial machine learning? Be specific

④ What are the contemporary issues of financial machine learning?

⑤ Why do financial machine learning projects usually fail?