ARTIFICIAL INTELLIGENCE IN INVESTMENT MANAGEMENT

STUDY GROUP 2021-2022, PROGRAMMING/CONTINUING EDUCATION COMMITTEE, CFA SOCIETY JAPAN

[A] General Sessions

- **Purposes:** Participants are expected to [1] learn basics of artificial intelligence (especially machine learning) in investment management and then [2] earn PL credits from CFA Institute. Educational materials by CFA Institute are presented in the **References** section below. We assume some (or all) of the participants move onto the next [B] Special Sessions after completion of this [A].
- **How we proceed:** Participants study by themselves by using the materials and ask questions on monthly Zoom calls. Each Zoom call covers a certain pre-specified part of the materials. Moderators and other particpants answer to the questions. Additionally, we encourage participants to catch up on a voluntarily and (bi-)weekly basis.
- **Deliverable(s):** N/A. (Each participant will earn PL credits from CFA Institute as a personal achievement.)
- **Moderator(s):** Akio Sashida, CFA and Yoshimasa Satoh, CFA + other volunteers

[A] References

- Robert Kissell, PhD, and Barbara J. Mack. 2019. "Fintech in Investment Management." CFA Institute Refresher Reading 2022 CFA Program Level I Reading 55. https://www.cfainstitute.org/membership/professional-development/refresher-readings/fintech-investment-management. (*)
- Kathleen DeRose, CFA, and Christophe Le Lannou. 2021. "Machine Learning." CFA Institute Refresher Reading 2022 CFA Program Level II Reading 4. https://www.cfainstitute.org/membership/professional-development/refresher-readings/machine-learning. (*)
- (Optional) Söhnke M. Bartram, Jürgen Branke, Mehrshad Motahari. 2020. "Artificial Intelligence in Asset Management." Research Foundation Literature Reviews https://www.cfainstitute.org/en/research/foundation/2020/rflr-artificial-intelligence-in-asset-management. (*)
- (Optional) Yoshimasa Satoh, CFA. 2021. "Introduction of Big Data and AI." https://github.com/yoshisatoh/ML/blob/main/References/Intro_of_BD_and_AI_2021-10-24.pdf

* These documents are only for CFA Institute members and please do not share them with other non-members.

[B] Special Sessions

- **Purposes:** Participants with basic / advanced knowledge and experience in artificial intelligence are expected to complete a research and developement project together. Research themes should be practically and/or academically meaningful.
- **How we proceed:** First off, participants discuss and decide what kind of project they hope to launch. For guidance, please refer to Exhibit 37 Stylized Decision Flowchart for Choosing ML Algorithms of "Machine Learning" which is introduced in the **References** section. Then the project will be managed in a democratic and collaborative manner. Tasks and progress will be reviewed on regular monthly Zoom calls.
- **Deliverable(s):** [B1] co-author and publish a blog article at https://blogs.cfainstitute.org/investor/ or a research paper at https://www.cfainstitute.org/en/research/financial-analysts-journal, [B2] make a presentation at a seminar/webinar by CFA Society Japan, and/or [B3] co-develop Python programs to analyze data in a machine learning approach and gain insights from that data (Note: Deliverables will be intelectural property in co-ownership amongsts direct contributors. It can be publicly presented on each individual's personal page of Github https://github.com/, based on mutual agreements.)
- **Moderator(s):** Akio Sashida, CFA and Yoshimasa Satoh, CFA + other volunteers

[A] [B] Schedules

Tentative schedules, which might vary, are as follows.

Table 1 – Timeline of [A] General Sessions

Sessions	Name	Pre-assignment (See References)	Monthly Zoom call (JST)
[A] General	*Intro*	*N/A*	• *2021/9/22 (Wed) 18:30 [Completed]* • *2021/9/28 (Tue) 11:30 [Completed]*
[A] General	*Fintech in Investment Management*	• *"Fintech in Investment Management." (p.1-25)*	• *2021/10/20 (Wed) 19:00 [Completed]*
[A] General	**Machine Learning #1**	• **"Machine Learning." (p.1-35, from 1. Introduction to 8. Ensemble Learning and Random Forest)**	• **2021/11/17 (Wed) 19:00 https://bit.ly/3aUoTIl (at the beginning of the call)**
[A] General	Machine Learning #2	• "Machine Learning." (p.35-73, from 9. Unsupervised Machine Learning Algorithms to the end)	• 2021/12/15 (Wed) 19:00 (at the beginning of the call)

Table 2 – Timeline of [B] Special Sessions

Sessions	Name	Pre-assignment (See References)	Monthly Zoom call
[B] Special	• **1. Define a problem to be solved.**	• **Please answer to the following questionnaire by 17th Nov 2021.** • **http://www.polljunkie.com/poll/szgfjg/study-group-2021-2022-cfa-society-japan**	Monthly review calls on the 3rd Wed of the month: • **2021/11/17 (Wed) 19:00 https://bit.ly/3aUoTIl**

Sessions	Name	Pre-assignment (See References)	Monthly Zoom call
	• **2. Prepare (alternative/big) data and and pre-process it by cleaning and standardizing.**		
[B] Special	• Ditto	• TBD	• 2021/12/15 (Wed) 19:00
[B] Special	• **3.** Evaluate machine learning algorithms including regularization.	• TBD	• 2022/1/19 (Wed) 19:00 • 2022/2/16 (Wed) 19:00 • 2022/3/16 (Wed) 19:00
[B] Special	• **4.** Improve the outcome by reviewing 2 and 3.	• TBD	• 2022/4/20 (Wed) 19:00 • 2022/5/18 (Wed) 19:00 • 2022/6/15 (Wed) 19:00 • 2022/7/20 (Wed) 19:00 • 2022/8/17 (Wed) 19:00 • 2022/9/21 (Wed) 19:00
[B] Special	• **5.** Present final results.	• TBD	• 2022/10/19 (Wed) 19:00

[B] 1. Define a problem to be solved.

Potential problems to be solved through this R&D initiative, are as follows.

Table 3 – Problems to be solved

#	Purpose	R&D theme	Data/Analysis	Potential data sources and providers (See the next section)
1A	Data Pre-processing	• **Consumer Sentiment Analysis**	• Web scraping / traffic • Social Media	• Publicly available data on the internet • FactSet
1B	Data Pre-processing	• **Corporate Sentiment Analysis**	• Quarterly / annual reports	• Publicly available
1C	Data Pre-processing	• **ESG Scoring**	• Quarterly / annual reports	• Publicly available

#	Purpose	R&D theme	Data/Analysis	Potential data sources and providers (See the next section)
2A	Alpha Signals & Market Risk Mgmt	• **Portfolio Mgmt**	• Traditional financial market data (e.g., equities, rates, credit spreads, currencies, etc.)	• Bloomberg, FactSet, Refinitiv, Nasdaq, etc.
2B	Alpha Signals & Market Risk Mgmt	• **Asset Allocation**	• Traditional macro-economic data	• Ditto
2C	Alpha Signals & Market Risk Mgmt	• **Finding non-traditional alpha and risk signals**	• Non-traditional assets (e.g., bitcoin price and volume)	• Ditto
3	Investment Consulting	• **Robo-advisor, pension consulting support tool**	• Traditional market data and non-traditional portfolio construction	• Ditto
4	Others	• (Please specify.)	•	•

[B] 2. Prepare (alternative/big) data: evaluation of alternative data providers

Alternative data provideds and evaluation results are as follows.

Table 4 – Evaluation of Alternative Data Providers

#	Name	Data Coverage and Venues	Free-trial / sample data	One-off/recurring fees in JPY terms	Availability on publications
1	FactSet	• https://www.factset.com/marketplace/catalog • FactSet, AWS, and Snowflake	• TBC	• TBC	• TBC
2	JADAA (**)	• https://alternativedata.or.jp/en/ • Platformers (BBG, FactSet, Refinitiv) or data source-specific	• TBC	• TBC	• TBC
3	Neudata	• TBD	• TBC	• TBC	• TBC
4	Nasdaq	• https://data.nasdaq.com/search • Python, R, or raw data (in jason, xml, csv)	• Yes (*)	• Free or TBC	• Yes
5	SEC EDGAR Company Filings	• https://www.sec.gov/edgar/sec-api-documentation • 10-K, 10-Q, etc. • Python (to get jason and txt files)	• Yes	• Free	• Yes

* In addition to sample data for their Premium Data, some data (e.g., Federal Reserve Economic Data) are available for free of charge. ** Japan Alternative Data Accelerator Association

Zoom Calls:
#1 2021/10/12 (Tue) 11:00 JST
#2 2021/10/20 (Wed) 11:00 JST
#3 2021/10/26 (Tue) 21:30 JST https://bit.ly/302HFLm
#4 N/A, but Yoshimasa Satoh knows how things work and can elaborate on it.
#5 Ditto.

[B] 3. Evaluate machine learning algorithms: selection of a machine learning algorithm

Potential machine learning algorithms are as follows.

Table 5 – Selection of a Machine Learning Algorithm

#	Input Data	Learning	Target	Target (detail)	ML method (*)
1	Numerical	• Supervised	• Classification	• Binary-class (0 or 1) • Multi-class	• Logistic • SVM • KNN • Naive Bayes • DNN
2	Numerical	• Supervised	• Regression	• Value	• (MLR) • DNN
3	Numerical	• Unsupervised	• Classification	• Binary-class (0 or 1) • Multi-class	• K-Means
4	Numerical	• Unsupervised	• Dimension Reduction	• PCA1, 2, …	• PCA
5	Text	• Supervised	• Classification (Sentiment Analysis)	• Multi-class (e.g., positive, negative, uncertainty, etc.)	• NLP

* Many of these methods can be both linear and non-linear. MLR (Multiple Linear Regression) and PCA (Principal Component Analysis) are linear only. Technically, MLR is a traditional statistical technique, but not a machine learning method.

[A] [B] Members and Moderators (indicated by boldface)

- **Akio Sashida, CFA a.sashida80@gmail.com [A][B]**
- Erik David McBain, CFA erik.mcbain@boosted.ai [B] (Ottawa, Ontario, Canada)
- Hidehito Tsujii htsujii@hotmail.com [B]
- James Sullivan, CFA j.sullivan@refinitiv.com [B]
- Ka You yohkax@icloud.com [B]
- Masayuki Teraguchi, CFA masateraguchi@yahoo.co.jp [B]
- Michinori Kanokogi, CFA mkanokogi@gmail.com, kanokogi_m867@nam.co.jp [A][B]
- Naoya Sugimoto, CFA naoya.sugimoto@gmail.com [B]

- Piotr Pomorski, CFA piotr.pomorski@outlook.com [B] (London, UK)
- Ranran Lin, CFA linranran27@hotmail.com [A][B]
- Shuxin Yang, CFA shuxin.yang@toki.waseda.jp [B]
- Tomokuni Higano, CFA tomokuni.higano@gmail.com [B]
- Tsuyoshi Yamakawa, CFA fifth.of.may0505@i.softbank.jp [A]
- Yasuhiro Oshima, CFA yasuhiro11oshima@gmail.com [A]
- **Yoshimasa Satoh, CFA satoh@cfaj.org [A][B]**

*Some of people cannot be found on Member Directory Search (https://directory.cfainstitute.org/). If you are a CFA charterholder, please let Yoshimasa Satoh know.

NOTE: This document will be updated on a regular basis as this iniative goes along.