Columbia | Engineering
The Fu Foundation School of Engineering and Applied Science
Contact information: JIFENG.LI
: jl6962@Columbia.edu,
: +13478037567,
WeChat: q2464498334)
Columbia University
IN THE CITY OF NEW YORK

Orchestration Framework for Financial Agents: From Algorithmic Trading to Agentic Trading

Jifeng Li, Arnav Grover, Abraham Alpuerto, Yupeng Cao, Xiao-Yang Liu

STEVENS INSTITUTE OF TECHNOLOGY 1870

Github Code

NEURAL INFORMATION PROCESSING SYSTEMS

Background

Classical **Algorithmic Trading (AT)** uses a fixed pipeline: **data → signals → risk → portfolio → execution → evaluation**. Building and maintaining such systems usually needs a full quant team and years of tuning.

Recent **LLM agents** can reason, call tools, and remember states, but most open-source trading "agent" projects are still ad-hoc and lack a standard orchestration layer or safety rules.

Financial markets are **mission-critical**: timestamps are strict, signals are weak, and data leakage can easily break any backtest.

Motivation

We want to turn a traditional AT system into an agentic system, without losing professional standards on risk, backtesting, and audit.

The framework should:

- reuse the same pipeline across stocks and crypto;
- separate LLM reasoning from numerical tools, to reduce data leakage and meta-overfitting;
- keep a memory of signals, prompts, and decisions for later audit and reuse.

The factor–asset graph (e.g., AAPL and NVDA connected to momentum/volatility/reversion factors) illustrates how signals are organized and shared across agents instead of being hard-wired in code.

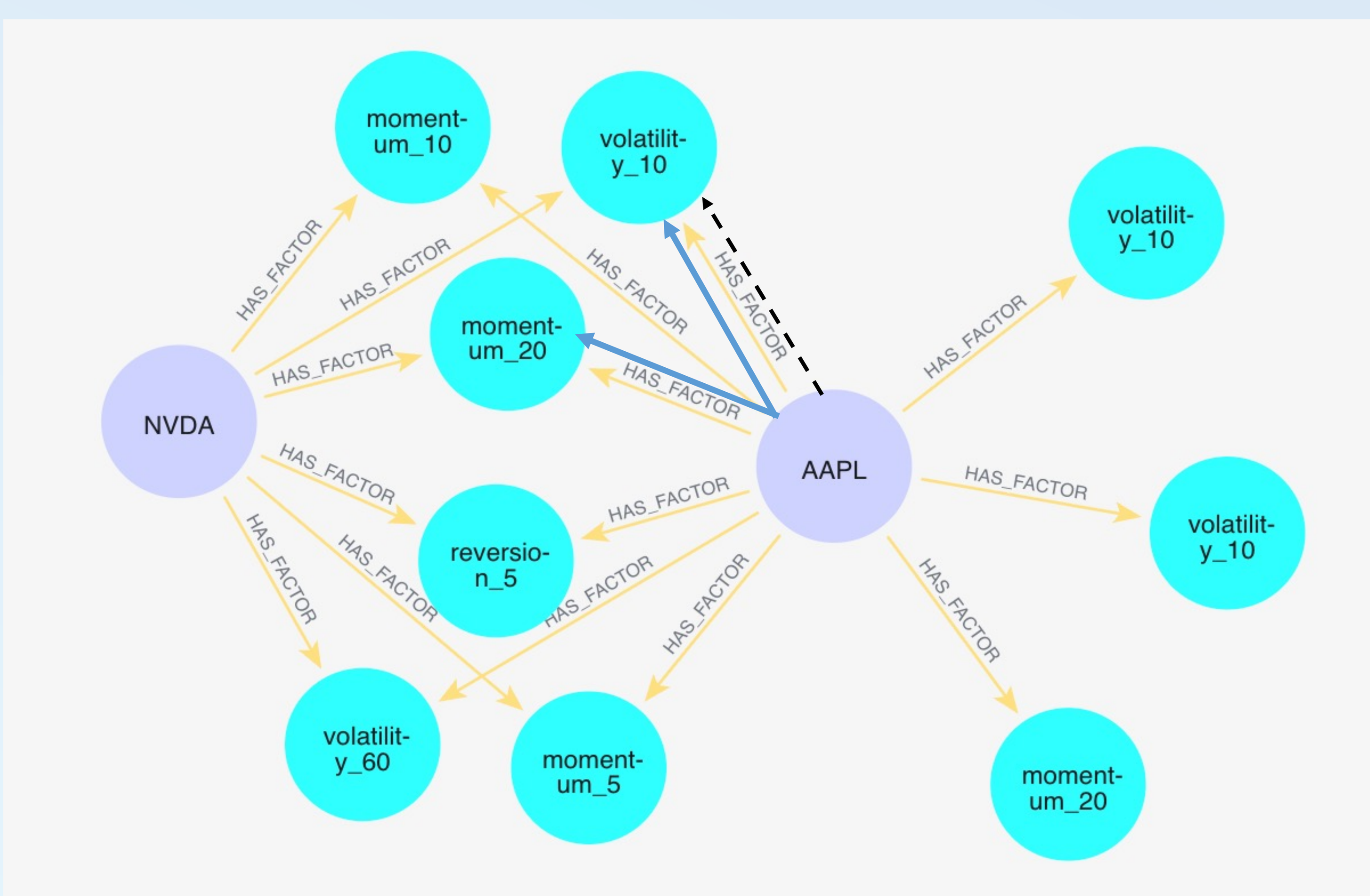

Figure 2: While RAG only retrieves the direct match (`volatility_10, dashed arrow`), the graph engine traverses the `HAS_FACTOR` edges (blue arrows) through the central asset (`AAPL`) to uncover hidden dependencies, such as the causal `momentum_20` factor.

Method (Agentic Trading)

Agent pools for each AT block
Planner/Orchestrator, Data, Alpha, Risk, Portfolio, Execution, Backtest, and Memory pools mirror a classic trading pipeline but are controlled by one DAG-style planner.

Protocols & safety
MCP handles structured control messages; A2A handles agent-to-agent chat. Evaluation-window prices and P&L never enter LLM contexts—only tools see labels; memory stores UUID-indexed summaries only.

Factor & signal design
Alpha Agents read training-window summaries and literature priors, output JSON factor specs (name, family, horizon, expected direction); numerical tools compute signals and portfolios outside the LLM.

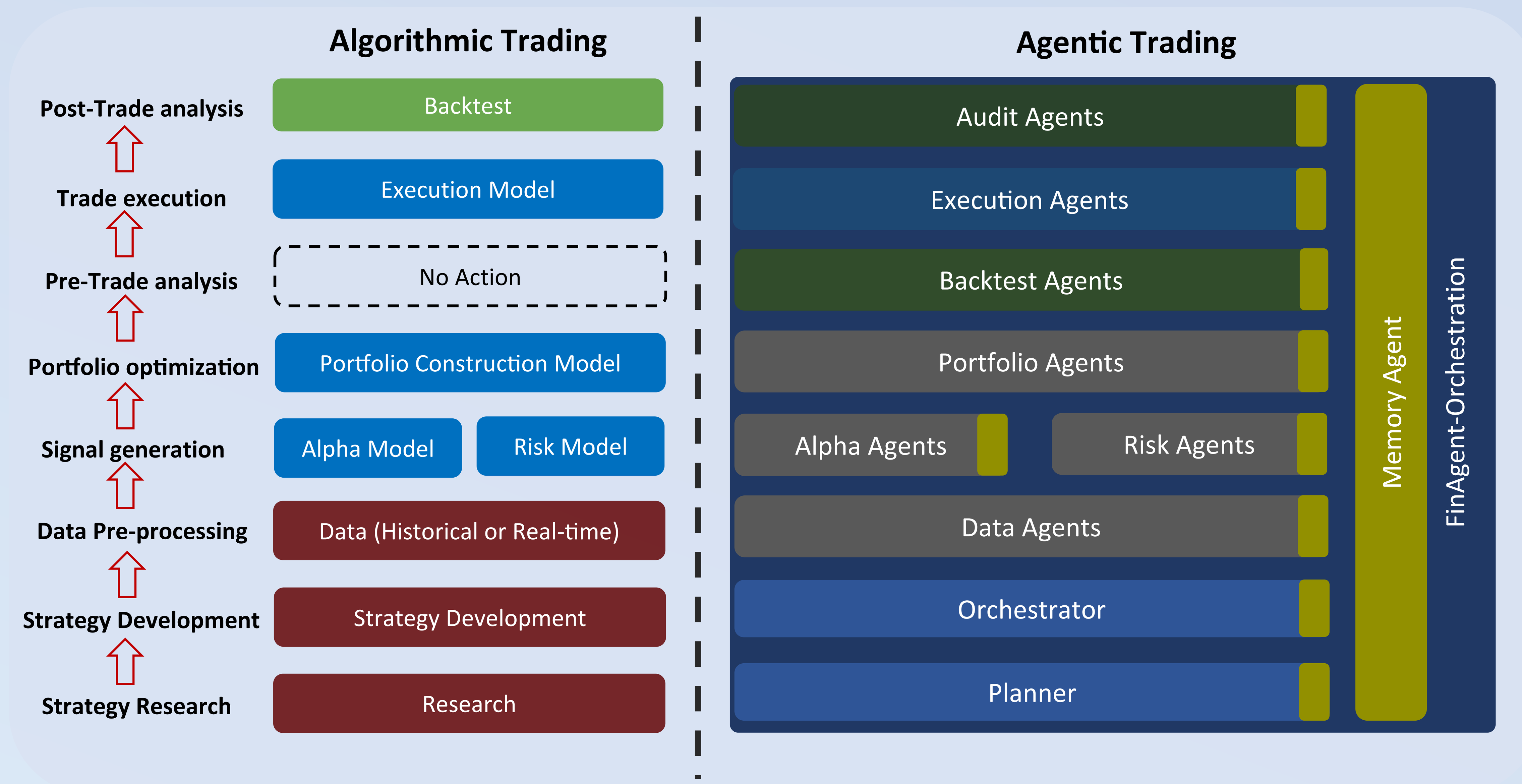

Figure 1: Agentic trading vs. algorithmic trading: we map the AT components to agents in our FinAgent orchestration framework, where a memory agent provides the contexts to other agents

Metric	Ours	SPY (B&H)	QQQ (B&H)	IWM (B&H)	VTI (B&H)	EW (Equal-Weighted)
Total Return (↑)	20.42%	16.60%	21.59%	11.45%	16.29%	**47.46%**
Annual Return (↑)	31.08%	25.07%	32.94%	17.10%	24.59%	**76.07%**
Volatility (↓)	**11.83%**	13.49%	18.38%	21.61%	13.72%	22.54%
Sharpe Ratio (↑)	2.63	1.86	1.79	0.79	1.79	**3.37**
MDD (%) (↑, less negative is better)	**-3.59**	-8.89	-14.13	-11.6	-9.06	-16.21

Table 1. Stocks – Trading Performance (R_f = 0; ↑ higher is better; ↓ lower is better; MDD = Maximum Drawdown; B&H = Buy & Hold; EW = Equal-Weighted Portfolio)

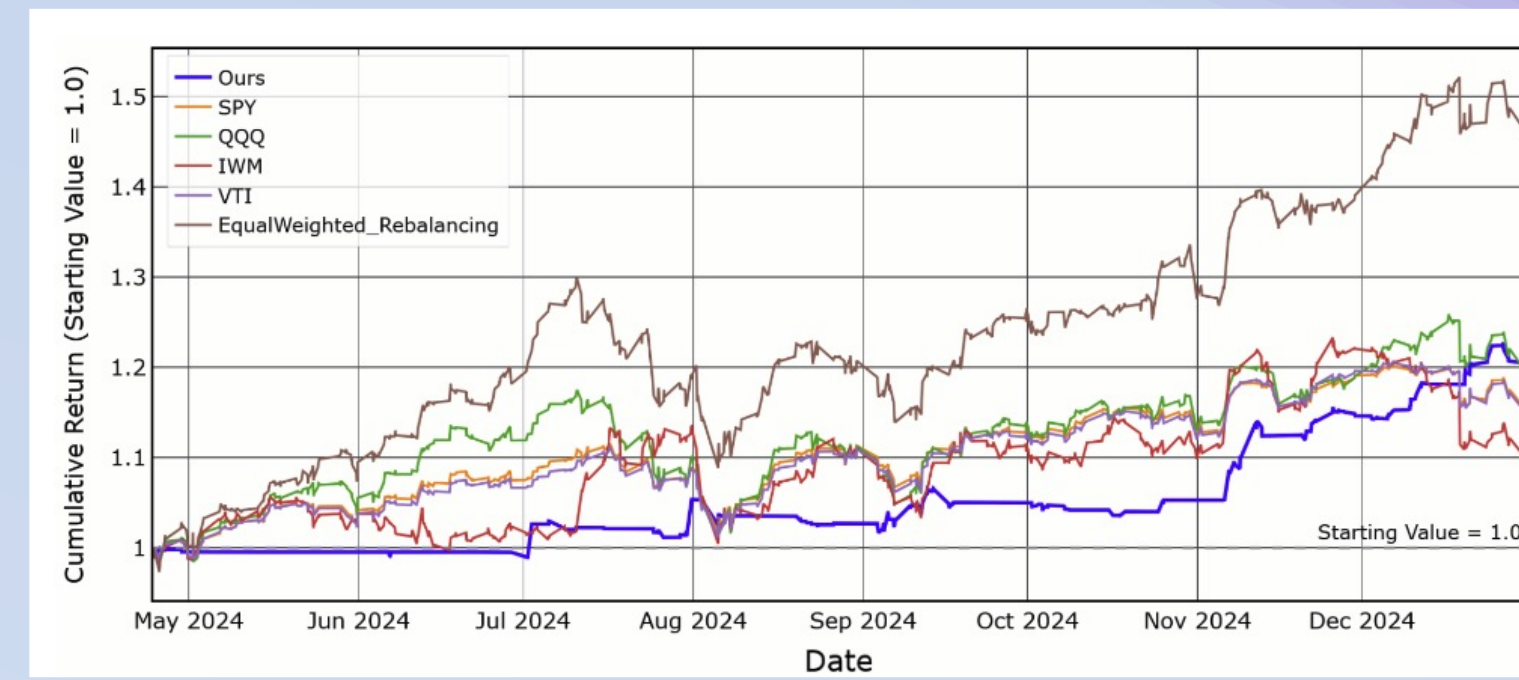

Fig 3. US Equities (Apr–Dec 2024): Agentic Strategy vs ETF & Equal-Weighted Benchmarks Seven-stock cumulative returns with a 3-month rolling training window. Our agent shows lower volatility and smaller max drawdown, while the equal-weighted portfolio achieves the highest total return (ETF baselines: SPY, QQQ, IWM, VTI; metrics in Table 1).

Test Results

US equities (hourly, Apr–Dec 2024, 7 stocks)

- Baselines: SPY, QQQ, IWM, VTI, and an equally-weighted (EW) portfolio with weekly rebalancing.
- Our agentic strategy: **20.42%** total return, **11.83%** volatility (lowest), Sharpe **2.63**, max drawdown **−3.59%** (smallest).
- EW: highest return **47.46%** but much higher volatility (**22.54%**) and drawdown (**−16.21%**).

BTC/USDT (minute, Jul 27–Aug 13 2025)

•Baseline: Buy-and-Hold; rolling training window 7 days.
•Our strategy: **+8.39%** vs Buy-and-Hold **+3.80%**, excess **+4.59%**.
•Volatility: **24.23%** vs **25.82%**; max drawdown **−2.80%** vs **−5.26%**.

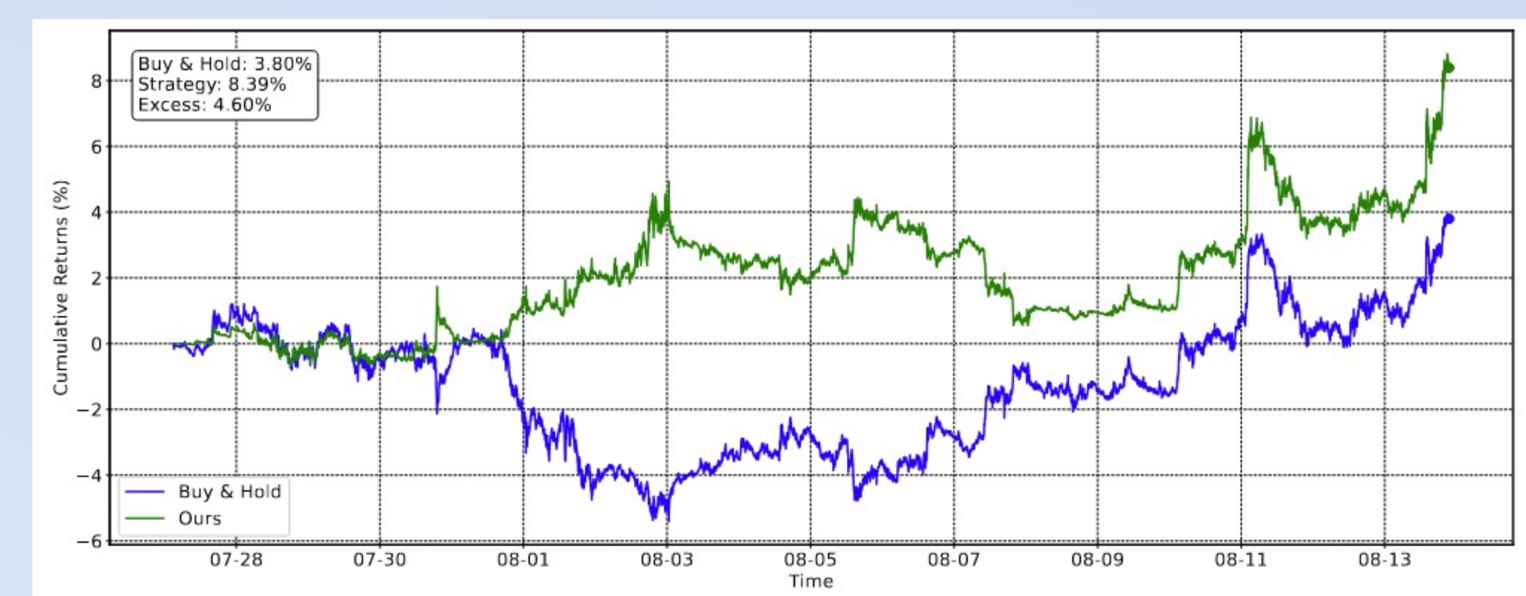

Fig 4. BTC Strategy vs Buy-and-Hold (Jul 27–Aug 13, 2025): Cumulative Returns. Cumulative returns: Buy-and-Hold +3.80%, Ours +8.39%, Excess +4.59%. Scrolling training window = 7 days; Excess = Ours − Buy-and-Hold.

Metric	BTC (Ours)	BTC (B&H)
Total Return (↑)	**8.39%**	3.80%
Annual Return (↑)	–	–
Volatility (↓)	**24.23%**	25.82%
Sharpe Ratio (↑)	**0.378**	0.17
MDD (%) (↑, less negative is better)	**-2.8**	-5.26

Table 2. BTC – Trading Performance (R_f = 0; ↑ higher is better; ↓ lower is better; MDD = Maximum Drawdown; B&H = Buy & Hold)

Conclusion

- Agentic Trading turns a standard algorithmic trading stack into **coordinated agent pools** with a single planner.
- **Leakage-aware protocols** and a Memory Agent allow reuse and audit of agent behavior without exposing evaluation labels.
- Experiments on US equities and BTC show **risk-controlled performance** and positive excess returns under the same orchestration pipeline.

Reference
[1] J. Li, A. Grover, A. Alpuerto, Y. Cao, and X.-Y. Liu, "Orchestration Framework for Financial Agents: From Algorithmic Trading to Agentic Trading," NeurIPS 2025 Workshop on Generative AI in Finance, 2025.
[2] Philip Treleaven, Michal Galas, and Vidhi Lalchand. Algorithmic trading review. Communications of the ACM, 56(11):76–85, 2013.
[3] Xiao-Yang Liu, Jingyang Rui, Jiechao Gao, Liuqing Yang, Hongyang Yang, Christina Dan Wang, Jian Guo, and Zhaoran Wang. FinRL-Meta: A universe of near-real market environments for data-driven deep reinforcement learning in quantitative finance. In NeurIPS 2021 Workshop on Data-Centric AI, 2022.