Enhancing fund recommendations with multi-round feedback-based reinforcement learning

YanShen YI, Chao WANG, Ying SUN, Qi ZHANG, Xunpeng HUANG, Hui XIONG

Frontiers of Computer Science, DOI: 10.1007/s11704-026-51887-2

Research Problem& Ideas

Why existing fund recommendation is hard, and how MAFR-RL responds.

Problems

Sparse and fragmented feedback

Most personalized recommenders still rely on collaborative filtering, which needs large-scale interaction logs. In practice, privacy regulations and platform fragmentation make those data difficult to obtain.

Heterogeneous fund attributes

Mutual funds have multi-view characteristics such as profitability, risk, and investment domains, yet prior methods often fail to encode them into one unified representation.

Fund–stock hierarchy is overlooked

A mutual fund is a portfolio of multiple stocks. Good recommendations should consider both fund-level properties and constituent-stock preferences, rather than only the fund as a single item.

Ideas

Interactive learning instead of static logs

Use a multi-round rating scenario with a simulated user to elicit preferences round by round and reduce reliance on historical interaction data.

Multi-view representation learning

Encode profitability, risk, investment domain, and top constituent stocks with specialized encoders plus self-attention to build a richer state for decision making.

Joint policy and value design

Adopt a hierarchical MoE actor, dual critics, and a tailored candidate-pool strategy so the policy can balance relevance, stock preference, and exploration.

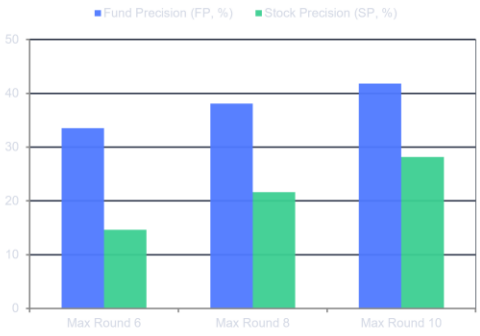
Research Problem& Key Contributions

Includes a results chart (FundData-2023) showing improvement with more rounds.

Key technical contributions

- Multi-round ratings-based RL setup + simulated-user protocol for sparse/fragmented logs.
- Multi-view fund representation (profitability, risk, investment domains) + top-K constituent stocks.
- Hierarchical MoE actor with Mamba experts for adaptive routing and fusion across views.
- Dual-critic objective (fund-level + stock-level) and candidate-pool sampling for robustness.

Results highlight (FundData-2023)



Both FP and SP improve as the max interaction rounds increase.

Numbers: 33.50/14.65 → 38.08/21.60 → 41.83/28.17 (FP/SP, %)