



UTokyo

Are Generative Al Agents Effective Personalized Financial Advisors?

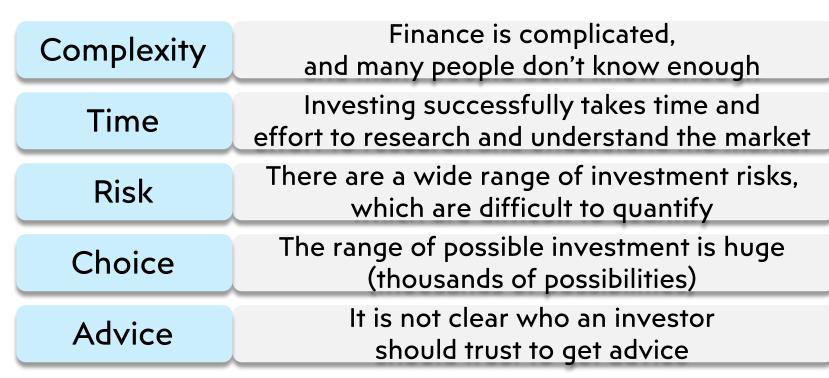
Takehiro Takayanagi, Kiyoshi Izumi

takayanagi-takehiro590@g.ecc.u-tokyo.ac.jp, izumi@sys.t.u-tokyo.ac.jp

Javier Sanz-Cruzado, Richard McCreadie, ladh Ounis {javier.sanz-cruzadopuig, richard.mccreadie, iadh.ounis}

@glasgow.ac.uk

Finding the right investments is hard!



Can we use AI/ML to help people to identify useful investments?

Conversational Generative Al agents can help!

Goal: Guide investors to choose the best financial assets through multiturn natural language conversations



Investor

Amazon's AWS segment drives significant revenue and growth, offering cloud services less impacted by economic swings....

to economic swings...

Agent

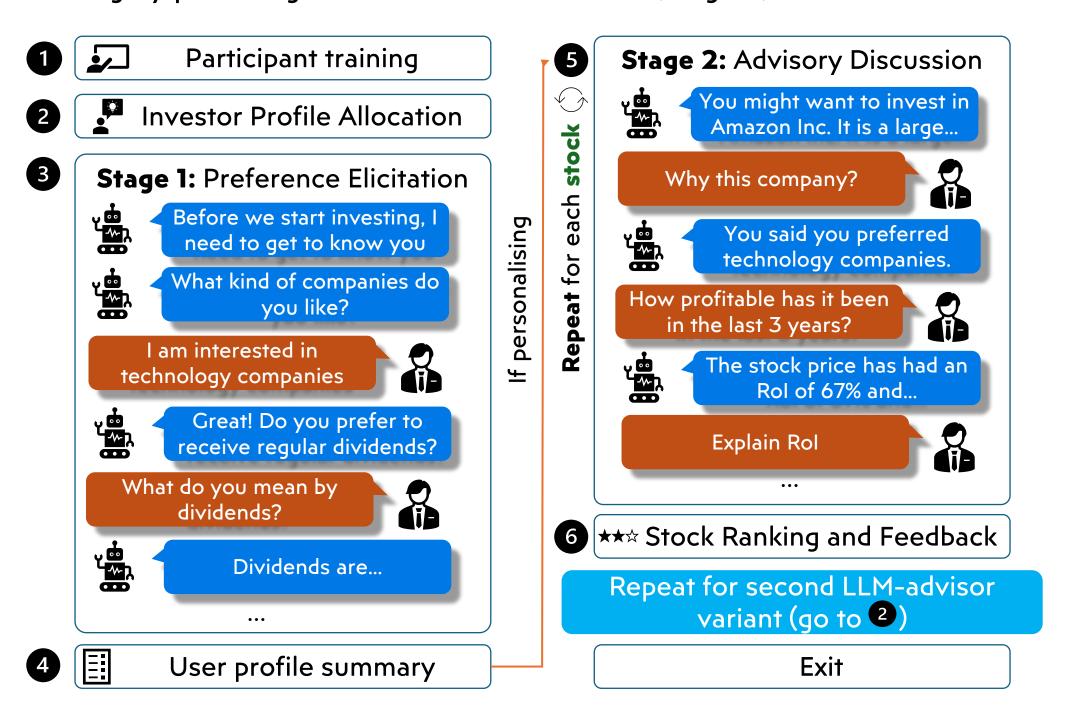
Agent

How good are LLMs at providing investment guidance?

We answer this through a user study!

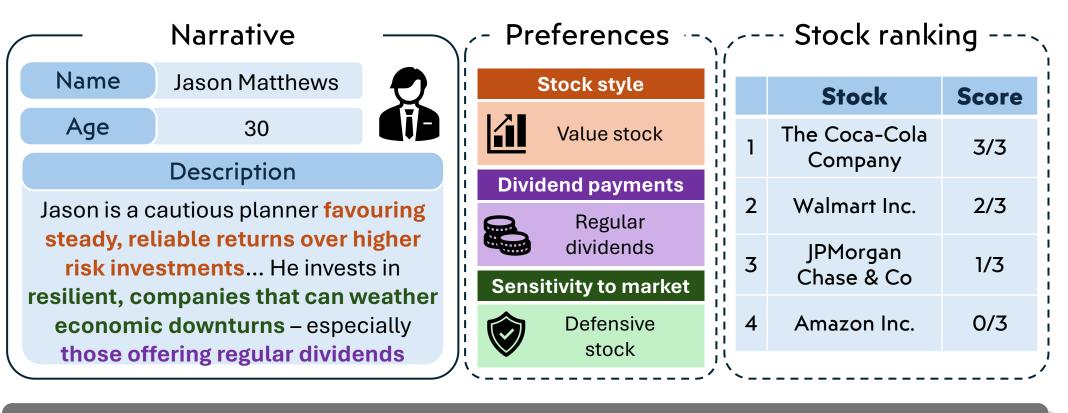
Design

- User task: Identify suitable stocks with the help of the LLM advisor and rank them by likelihood to buy.
- LLM advisor task: Obtain user preferences (stage 1) and support decisionmaking by providing information about stocks (stage 2).



Expert-designed Gold Standard

- **Problem:** how can we evaluate how good LLM agents are?
- Solution: Users role-play as pre-defined investor archetypes, designed in collaboration with financial experts.
- Expert-curated Gold Standards:
 - Investment preferences: evaluate preference elicitation (stage 1).
 - Ground truth stock rankings: evaluate user decision quality (stage 2).



Other information

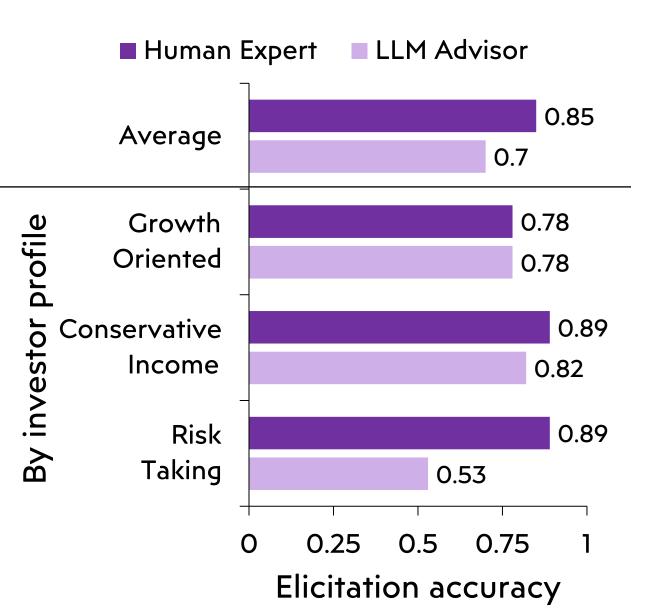
- Participants: 60 students from the UK and Japan.
- Each participant uses two different LLM advisor variants.
- 3 investor narratives: growth oriented, conservative income, risk taking.
- 4 different stocks for each investor narrative (picked by experts).

Preference elicitation

Stage 1

Can LLM advisors effectively obtain user preferences through conversation?

We compare the effectiveness of an LLM advisor against a human expert advisor



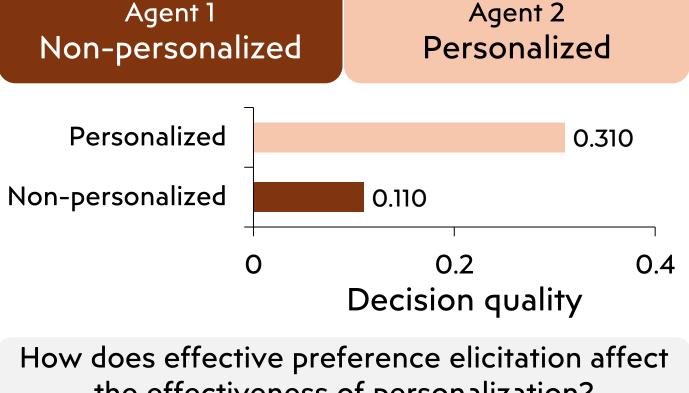
LLMs are promising for obtaining investor's preferences, but not yet robust across all user types.

Personalized guidance

Results

Stage 2, Experiment 1

Does personalization lead to better investment decisions?



the effectiveness of personalization?

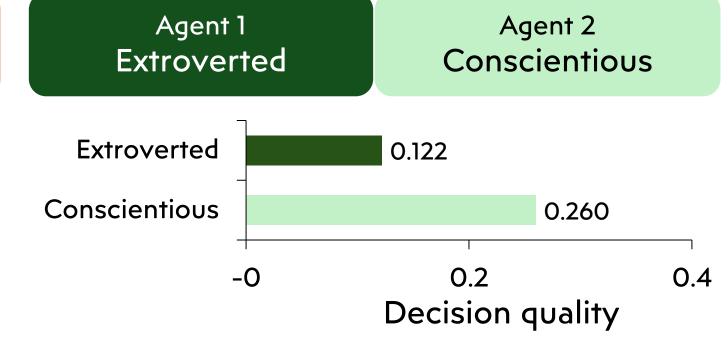


Agent personalization improves decisionmaking, but bad elicitation leads to harmful advice.

Personality and trust

Stage 2, Experiment 2

Does the personality of personalized advisors affect investor's decision quality and perception?



- Conscientious advisor leads to better decision quality.
- However, users preferred the extroverted advisor!
 - Intention to use: +11.88% more likely to use the extroverted advisor
 - Overall satisfaction: +5.88% greater satisfaction from the extroverted advisor

Users cannot distinguish between good and bad advice. Trust is driven by personality, not accuracy.