InvestAlign: Overcoming Data Scarcity in Aligning Large Language Models with Investor Decision-Making Processes Under Herd Behavior

Anonymous ACL submission

Abstract

Aligning Large Language Models (LLMs) with investor decision-making processes under herd behavior is a critical challenge in behavioral finance, which grapples with a fundamental limitation: the scarcity of real-user data needed for Supervised Fine-Tuning (SFT). While SFT can bridge the gap between LLM outputs and human behavioral patterns, its reliance on massive authentic data imposes substantial collection costs and privacy risks. We propose **InvestAlign**, a novel framework that constructs high-quality SFT datasets by leveraging theoretical solutions to similar and simpler optimal investment problems rather than complex scenarios. Our theoretical analysis demonstrates that training LLMs with **InvestAlign**-generated data achieves faster parameter convergence than using real-user data, suggesting superior learning efficiency. Furthermore, we develop **InvestAgent**, an LLM agent fine-tuned with **InvestAlign**, which demonstrates significantly closer alignment to real-user data than pre-SFT models in both simpler and complex investment problems. This highlights **InvestAlign** as a promising approach with the potential to address complex optimal investment problems and align LLMs with investor decision-making processes under herd behavior.

1 Introduction

In financial markets, investors typically make decisions based on their risk preferences to achieve higher returns, lower volatility, and maximize their utility. (Merton, 1969). Investment decisions are crucial as they not only impact individual financial outcomes but also shape market dynamics and overall economic stability, making them a key driver of both personal wealth and broader market efficiency (Ahmad and Wu, 2022). During this process, investment assistants such as financial analysts and fund managers, play a significant role by sharing their own investment decisions through platforms (Brown et al., 2008). These investment assistants often have rich investment experience and extensive influence, leading investors to mimic their behaviors. This is commonly referred to as herd behavior in microeconomics and behavioral finance (Bikhchandani and Sharma, 2000). The prior works in (Wang and Zhao, 2024a,b, 2025) have investigated the optimal investment problem considering herd behaviors between two agents, and theoretically analyzed the impact of herd behavior on their decisions. However, there are more complex problems where the above models fall short or only provide qualitative insights (Zhou and Liu, 2022), prompting us to explore alternative approaches.

Large Language Models (LLMs) have been widely adopted in various domains as generative agents to assist with specific tasks (Kovač et al., 2023; M. Bran et al., 2024). A notable trend is the enhancement of LLM agents with human-like intelligence to simulate human decision-making processes (Gao et al., 2024). In economics and finance, substantial works have been done on aligning LLMs with human values and decisions, particularly in models for market behavior prediction and the analysis of complex economic data for policy-making (Zhao et al., 2023; Lee et al., 2024). These studies predominantly address macroeconomic concerns, such as the dynamics of information dissemination and collective decision-making within global markets (Li et al., 2024b). To our best knowledge, there has been limited exploration of LLMs' efficacy in microeconomics and behavioral finance, and current LLMs do not fully align with investor decision-making processes, as shown in Section 3.

Achieving the alignment of LLMs to investor decision-making processes often relies on large-scale real-user data in Supervised Fine-Tuning (SFT) (Zhang et al., 2023). Fine-tuned with specific training datasets, LLMs can better generate investor behavior in complex problems. However, it faces the following obstacles. Collecting real-

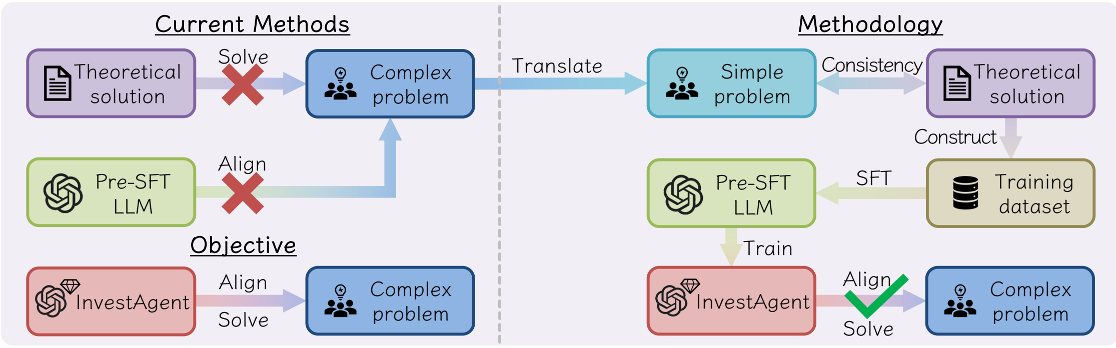


Figure 1: Overview of **InvestAlign**.

user data can be costly due to the wide variation in investment attributes like risk preference and herd degree (Abbot, 2017). Additionally, many investors are reluctant to share their investment decisions due to privacy and security concerns.

To address data scarcity, note that for some simple problems such as the one in (Wang and Zhao, 2024b), the theoretical solution has already been found, using which we can generate a large amount of training data. One possible solution is that, given a complex problem, we first identify a similar and simpler problem with a theoretical solution, construct the SFT dataset using this theoretical solution, and then fine-tune LLMs to solve the original complex problem. There are several issues to be addressed when following this approach:

- Q_1: Given the complex problem, how to identify a similar and simpler problem?
- Q_2: Do the theoretical solution of the simpler problem align with real users' investment decisions, and can they be used to construct a training dataset that mirrors investor decision-making processes?
- Q_3: How can we generate the training dataset based on the theoretical solution of the simpler problem? How does it align with investor decision-making processes compared with real-user data?
- Q_4: How to adapt the fine-tuned LLMs to solve the complex problem, and what is its performance?

To validate the feasibility of the proposed approach and address the above four issues, in this work, we examine an optimal investment scenario involving two agents as a case study. We consider the following two primary factors influencing herd behavior. The first factor is the pattern of herd behavior, which includes *absolute herd behavior* in (Wang and Zhao, 2024b), where agents replicate the entire portfolio of others, and *relative herd behavior* in (Wang and Zhao, 2024a), where agents mimic the changing rate of others' decisions. The second factor is the structure of the influence network structure, which includes *unilateral influence* from one agent to another and *mutual influence* between two agents (Wang and Zhao, 2025). We investigate two complex problems corresponding to relative herd behavior under unilateral influence and absolute herd behavior under mutual influence, respectively. Although theoretical solutions for these problems exist, their computational complexity is notably high. To address Q_1, we utilize absolute herd behavior under unilateral influence as the simpler problem, for which the theoretical solution is more readily derived. Note that while the simpler problem shares mathematical similarities with the original complex problems, they differ in their approaches to measuring herd behavior.

To answer Q_2, we collect real-user data on the simpler problem and apply statistical methods to validate the consistency between real-user data and the theoretical solution. Next, to answer Q_3, we construct SFT datasets based on the theoretical solutions, and theoretically prove that fine-tuning LLMs on the above training datasets leads to faster parameter convergence than using real-user data. Then, to answer Q_4, given the training dataset, we fine-tune the LLMs and develop the **InvestAgent**s, which can make decisions similar to the theoretical solution, thus aligning with real-user data in the simpler problem. Finally, we conduct another real-user test to verify the performance of **InvestAgent**s on solving the original complex problems, and experimental results show that **InvestAgent**s exhibit better alignment performance than pre-SFT LLMs.

In conclusion, our contributions include:

- We explore and utilize LLMs in micro economics and behavioral finance, particularly in the domain of optimal investment under herd behavior.
- We effectively construct a large amount of high-quality training data with the theoretical solution of the corresponding mathematical model.
- We propose the LLM alignment techniques, **InvestAlign**, using the generated abundant dataset and apply SFT to fine-tune LLMs.

2 Related Work

LLMs in Finance and Optimal Investment. For finance-related tasks, several specialized LLMs have been developed, e.g., BloombergGPT (Wu et al., 2023), FinGPT (Yang et al., 2023a), and XuanYuan (Zhang and Yang, 2023). The success of these models depends on large amounts of training data, and the challenge is how to effectively collect and generate high-quality data, which is a key goal of our proposed method. Focusing on the optimal investment problem, prior studies have explored the use of LLMs in different scenarios such as investment idea generation and quantitative investment (Li et al., 2023; Wang et al., 2023a). However, within agent-based modeling, only a few works use LLMs as generative investors to simulate or complement human investor behavior, e.g., InvestLM in (Yang et al., 2023b) and EconAgent in (Li et al., 2024b). Similar agent-based ideas have been widely used in many areas such as problems in the economic system (Horton, 2023; Chen et al., 2023; Geerling et al., 2023), social science (Ghaffarzadegan et al., 2023; Liu et al., 2024; Wang et al., 2024b), and natural science (Boiko et al., 2023; M. Bran et al., 2024). While several studies in other domains have explored the LLMs' irrational behaviors to mirror human cognitive biases (Liu et al., 2024; Wang et al., 2024a; Xiao et al., 2024), existing agent-based LLMs for investment have not yet accounted for the herd behavior (Bikhchandani and Sharma, 2000), which is significant in microeconomics and behavioral finance. Understanding its influence on the optimal investment problem while incorporating LLMs is crucial for analyzing investor behavior (Ahmad and Wu, 2022).

LLM Alignment. LLM alignment with human values has emerged as a critical area of research (Wang et al., 2024d), aiming to make LLM agents behave in line with human intentions and values (Ji et al., 2023). Although LLMs excel in various tasks, issues like untruthful answers (Bang et al., 2023), sycophancy (Perez et al., 2022), and deception (Steinhardt, 2023), raise concerns about controllability and risks in LLM agents. To achieve forward alignment, which ensures that trained systems meet alignment requirements, numerous methods for policy learning and scalable oversight are proposed (Ji et al., 2023; Wang et al., 2024d). For LLMs, a typical approach is reinforcement learning from human feedback (RLHF) using SFT (Christiano et al., 2017; Bai et al., 2022; Bowman et al., 2022; Wang et al., 2024c). In microeconomics and behavioral finance, only a limited number of studies involve LLMs (Li et al., 2024b; Horton, 2023), focusing on macro-level alignment while ignoring microcosmic behaviors of human decision-making.

SFT Methods in Optimal Investment. SFT is a widely adopted technique in the field of LLMs for improving model performance on specific tasks by refining pre-trained models with a dataset tailored to the target task (Zhang et al., 2023). Many tricks and methods of SFT have been proposed to achieve better LLM alignment to humans, e.g., (Ding et al., 2023; Wang et al., 2023b; Xie et al., 2024; Li et al., 2024a). In the domain of finance, SFT has been applied to various investment-related tasks such as price prediction, financial reports summarization, sentiment analysis, portfolio optimization, etc. (Zhao et al., 2024; Guo and Hauptmann, 2024; An et al., 2024). These advancements highlight the power of SFT in tailoring LLMs to meet the specific needs of investment strategies, enabling models to simulate or complement human-like behaviors. However, collecting large, high-quality datasets for fine-tuning in optimal investment remains a challenging problem (Abbot, 2017).

3 Problem Simplification & Real-User Data Verification

To verify the feasibility of the proposed method **InvestAlign**, we consider the optimal investment scenario involving two agents A_1 and A_2. As mentioned above, for the first complex problem P_1, we assume that A_1's investment decisions are unilaterally influenced by A_2 under relative herd behavior, and for the second problem P_2, we assume that A_1's and A_2's investment decisions are mutually influenced under absolute herd behavior. For the simple problem with the theoretical solution, denoted by P_3, we assume that A_1's investment decisions are unilaterally influenced by A_2 under absolute herd behavior. Next, to answer Q_2, we collect real-user data using interviews and questionnaires, and obtain pre-SFT LLMs' investment decisions for P_3. Then, we show that pre-SFT LLMs' responses are misaligned with the real-user data, and validate the statistical consistency between the theoretical solutions and the real-user data.

3.1 Optimal Investment Problems

Following the work in (Merton, 1969), we consider the scenario where A_1 and A_2 invest in the period

$\mathcal{T} = [0, T]$ in a financial market consisting of a deposit and a stock. We define A_i's fund invested in the stock as his/her *investment decisions*, denoted by $\{P_i(t)\}_{t\in\mathcal{T}}$ $(i = 1, 2)$. We denote r as the interest rate of the deposit, and v and σ as the excess return rate and volatility of the stock. Given the above parameters, A_i's fund $\{X_i(t)\}_{t\in\mathcal{T}}$ satisfies

$$\mathrm{d}X_i(t) = [rX_i(t) + vP_i(t)]\mathrm{d}t + \sigma P_i(t)\mathrm{d}W(t), \quad (1)$$

where $X_i(0) = x_{i,0}$ is his/her initial fund, and $\{W(t)\}_{t\in\mathcal{T}}$ is a standard Brownian motion modeling the randomness of the stock price. Considering the herd behavior, A_i jointly maximizes his/her expected utility of the terminal fund $\mathbb{E}\phi_i[X_i(T)]$ and minimizes the distance between his/her own and the other's decisions $D(P_1, P_2)$. Following the work in (Rogers, 2013), we assume that A_i's utility is $\phi_i[X_i(T)] = -\frac{1}{\alpha_i}\exp[-\alpha_i X_i(T)]$, where α_i is his/her risk aversion coefficient. In summary, the general optimal investment problem is

$$\begin{cases} \sup_{\{P_1(t)\}_{t\in\mathcal{T}}} \mathbb{E}\phi_1[X_1(T)] - \theta_1 D(P_1, P_2), \\ \sup_{\{P_2(t)\}_{t\in\mathcal{T}}} \mathbb{E}\phi_2[X_2(T)] - \theta_2 D(P_1, P_2), \end{cases} \quad (2)$$

where θ_i is A_i's influence coefficient to address the tradeoff between the two different objectives. We define the risk aversion coefficient α_i and the influence coefficient θ_i as A_i's *investment attribute*.

• P_1**: Optimal investment problem under relative herd behavior with unilateral influence.** Following the work in (Wang and Zhao, 2024a), when considering the relative herd behavior, the distance is defined as $\delta(P_1, P_2) = \frac{1}{2}\int_{\mathcal{T}}[P_1'(t) - P_2'(t)]^2\mathrm{d}t$, i.e., the integrated square error between the two decisions' changing rates, and when considering the unilateral influence of A_2 on A_1, A_2's influence coefficient $\theta_2 = 0$. In this case, the optimal investment problem (2) becomes P_1, which is

$$\begin{cases} \sup_{\{P_1(t)\}_{t\in\mathcal{T}}} \mathbb{E}\phi_1[X_1(T)] - \theta_1 \delta(P_1, P_2), \\ \sup_{\{P_2(t)\}_{t\in\mathcal{T}}} \mathbb{E}\phi_2[X_2(T)]. \end{cases} \quad (3)$$

• P_2**: Optimal investment problem under absolute herd behavior with mutual influence.** Following the work in (Wang and Zhao, 2025), when considering the absolute herd behavior, the distance is defined as $\Delta(P_1, P_2) = \frac{1}{2}\int_{\mathcal{T}}[P_1(t) - P_2(t)]^2\mathrm{d}t$, i.e., the integrated square error between the two agents' decisions, and when considering the mutual influence, the two agents' influence coefficients, θ_1 and θ_2, are both positive. In this case, the optimal investment problem (2) becomes P_2, which is

$$\begin{cases} \sup_{\{P_1(t)\}_{t\in\mathcal{T}}} \mathbb{E}\phi_1[X_1(T)] - \theta_1 \Delta(P_1, P_2), \\ \sup_{\{P_2(t)\}_{t\in\mathcal{T}}} \mathbb{E}\phi_2[X_2(T)] - \theta_2 \Delta(P_1, P_2). \end{cases} \quad (4)$$

• P_3**: Optimal investment problem under absolute herd behavior with unilateral influence** is

$$\begin{cases} \sup_{\{P_1(t)\}_{t\in\mathcal{T}}} \mathbb{E}\phi_1[X_1(T)] - \theta_1 \Delta(P_1, P_2), \\ \sup_{\{P_2(t)\}_{t\in\mathcal{T}}} \mathbb{E}\phi_2[X_2(T)]. \end{cases} \quad (5)$$

From the work in (Wang and Zhao, 2024b), A_i's theoretical optimal decision, denote by $\{\hat{P}_i(t)\}_{t\in\mathcal{T}}$, of P_3 can be easily calculated, as shown in (11) in Appendix A.1. We set the parameter values of P_1, P_2, and P_3 in Appendix A.2.

3.2 Data Collection

Real-User Data Collection. To verify whether P_3's theoretical solution in (11) matches users' investment decisions, we collect real-user data from 119 participants using interviews and questionnaires when facing the investment problem P_3. We denote the index set of participants as $\mathcal{I} = \{1, 2, \ldots, 119\}$. To reduce bias and noise in the collected data, we primarily recruit professionals and students in the fields of microeconomics and behavioral finance, and we treat the real-user data as a proxy for the ground truth. We let the participant play the role of A_1 unilaterally influenced by an investment assistant A_2 whose investment attribute is set in Appendix A.2.

The questionnaire we use is in Figure 15 in Appendix A.14. In the first part, we provide the task description, including the asset information and the participants' goals. In the second part, participants report their investment decisions, denoted by $\{\tilde{P}_1^i(t)\}_{t\in\mathcal{T}}$ for all $i \in \mathcal{I}$. To facilitate participants' decision-making, we ask them to report the proportions of funds invested in the stock $\{\tilde{P}_1^i(t)/X_1^i(t)\}_{t\in\mathcal{T}}$. When processing the data, we first calculate $\{X_1^i(t)\}_{t\in\mathcal{T}}$ using (1), and then calculate the participants' decisions $\{\tilde{P}_1^i(t)\}_{t\in\mathcal{T}}$ according to the proportions $\{\tilde{P}_1^i(t)/X_1^i(t)\}_{t\in\mathcal{T}}$. In the third part, we ask the participants the information about their investment attributes, based on which, we calculate their risk aversion coefficients $\{\alpha_1^i\}_{i\in\mathcal{I}}$ and influence coefficients $\{\theta_1^i\}_{i\in\mathcal{I}}$ as follows. Details are in Appendix A.3.

Collection of Pre-SFT LLMs' Investment Decisions. Next, to verify whether pre-SFT LLMs align with real-user data, we collect the pre-SFT LLMs' decisions. We choose a variety of LLMs, including API-based model `GPT-3.5-Turbo` (Achiam et al., 2023), as well as open-source models like `Qwen-2-7B-Instruct` (Yang et al., 2024) and `Meta-Llama-3.1-8B-Instruct` (Dubey et al., 2024). To obtain these pre-SFT LLMs' investment decisions in P_2, we first construct a prompt, as shown in Figure 11 in Appendix A.13. The first part is identical to the questionnaire in Figure 15, where we designate the pre-SFT LLM as an investment expert and describe the task. In the second part, we assign the pre-SFT LLM its investment attribute, corresponding to the participant's investment attribute $\{\alpha_1^i\}_{i\in\mathcal{I}}$ and $\{\theta_1^i\}_{i\in\mathcal{I}}$ in the real-user data. In the third part, the pre-SFT LLM reports the proportion of its funds invested in the stock $\{P_1^i(t)/X_1^i(t)\}_{t\in\mathcal{T}}$. We then obtain the pre-SFT LLM's investment decision $\{P_1^i(t)\}_{t\in\mathcal{T}}$.

3.3 Validation of Pre-SFT LLMs and the Theoretical Solution

The real-user data shows that the participants' risk aversion coefficients $\{\alpha_1^i\}_{i\in\mathcal{I}}$ and influence coefficients $\{\theta_1^i\}_{i\in\mathcal{I}}$ fall within the ranges of $\tilde{\mathcal{S}}_{\alpha_1} = [0.09, 0.38]$ and $\tilde{\mathcal{S}}_{\theta_1} = [0, 1\times 10^{-7}]$, respectively. For the convenience of data processing, we discretize these two sets into $\tilde{\mathcal{S}}_{\alpha_1} = \bigcup_{m\in\mathcal{M}} \tilde{\mathcal{S}}_{\alpha_1}^m$ and $\tilde{\mathcal{S}}_{\theta_1} = \bigcup_{n\in\mathcal{N}} \tilde{\mathcal{S}}_{\theta_1}^n$, and treat values that fall within the same interval as the same value[1]. We then group the participants according to these subsets, with participants sharing the same investment attributes forming a class. Specifically, the class of participants with risk aversion coefficient $\alpha_1 \in \tilde{\mathcal{S}}_{\alpha_1}^m$ and influence coefficient $\theta_1 \in \tilde{\mathcal{S}}_{\theta_1}^n$ for all $m \in \mathcal{M}$ and $n \in \mathcal{N}$ is denoted as $\mathcal{I}^{mn} = \{i|\alpha_1^i \in \tilde{\mathcal{S}}_{\alpha_1}^m, \theta_1^i \in \tilde{\mathcal{S}}_{\theta_1}^n\}$ for all $m \in \mathcal{M}$ and $n \in \mathcal{N}$.

For each participant class $\mathcal{I}^{mn}$, we calculate the mean and the 95% confidence interval of the real-user data, the mean and the 95% confidence interval of the pre-SFT LLMs' investment decisions based on 10 repeated trials with the same investment attribute, and the corresponding theoretical solution. Here, we take the investment attribute $\alpha_1 = 0.13$ and $\theta_1 = 7\times 10^{-8}$ as an example, and observe the same trend for other values. The experimental results are in Figure 2. In figure legends, we omit the subscript 1 from P_1 where no ambiguity arises. As shown in Figure 2, there is a significant discrepancy between the pre-SFT LLMs' investment decisions and the real-user data, indicating that pre-SFT LLMs fail to align with real-user data in optimal investment under absolute herd behavior. We also find that the performances of pre-SFT LLMs in P_1 and P_2 are misaligned, as shown in Appendix A.5, underscoring the necessity of supervised fine-tuning to bridge the gap between pre-SFT LLMs' decisions and real-user data.

On the contrary, from Figure 2, the theoretical solutions are much closer to the real-user data than pre-SFT LLMs' investment decisions. We further employ statistical methods to validate the consistency between the theoretical solutions and real-user data. For the i-th participant, we denote his/her real investment decision as $\{\tilde{P}_1^i(t)\}_{t\in\mathcal{T}}$, and the theoretical solution with the same investment attribute as $\{\hat{P}_1^i(t)\}_{t\in\mathcal{T}}$, respectively. We calculate the difference d_i and correlation coefficient ρ_i between $\{\tilde{P}_1^i(t)\}_{t\in\mathcal{T}}$ and $\{\hat{P}_1^i(t)\}_{t\in\mathcal{T}}$, and conduct t-tests on the means of the differences $\{d_i\}_{i\in\mathcal{I}}$ and the correlation coefficients $\{\rho_i\}_{i\in\mathcal{I}}$ (Shao, 2008), respectively, which show that there exists significant consistency between the theoretical solution and real-user data. Details are in Appendix A.6.

In summary, due to the significant gap between pre-SFT LLMs and real-user data, fine-tuning the LLMs with the theoretical solution is critical. As the theoretical solution closely aligns with real-user data, we can use them to construct the SFT dataset as a substitute for real-user data.

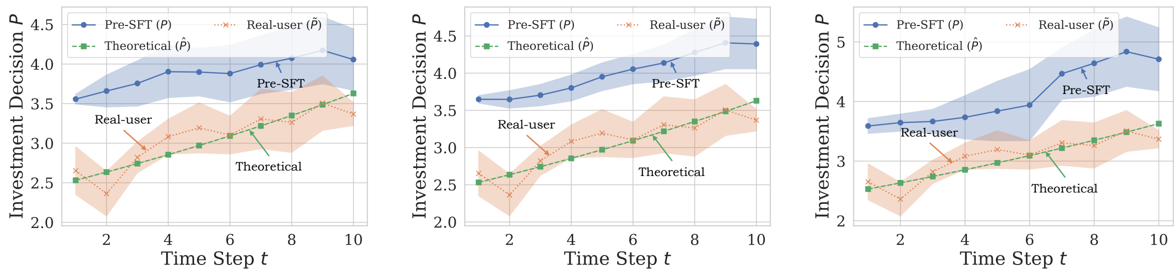
4 Methodology: InvestAlign

In this section, to answer Q_3, i.e., how we can construct the SFT dataset using the theoretical solution, and whether it performs better in fine-tuning compared to real-user data, we first introduce the method of constructing SFT datasets using the theoretical solution. Then, we theoretically prove that training LLMs on these datasets results in faster parameter convergence than using real-user data.

4.1 Constructing SFT Dataset with Theoretical Solution

The SFT dataset comprises input-output pairs for fine-tuning LLMs, which are generated based on a custom prompt template. The prompt for SFT is in Figure 12 in Appendix A.13. When constructing the SFT dataset, we need to vary A_1's investment attribute α_1 and θ_1. Following the work in (Wang and Zhao, 2024b), we set the values of

[1] Specifically, we set $\tilde{\mathcal{S}}_{\alpha_1} = [0.09, 0.13) \cup [0.13, 0.19) \cup [0.19, 0.26) \cup [0.26, 0.38) \cup \{0.38\}$ and $\tilde{\mathcal{S}}_{\theta_1} = \bigcup_{k=0}^{9} [k\times 10^{-8}, (k+1)\times 10^{-8}) \cup \{1\times 10^{-7}\}$, and use the left-point value to approximate the entire interval.



(a) `GPT-3.5` (b) `Qwen-2` (c) `Llama-3.1`

Figure 2: Comparison of real-user data ($\tilde{P}_1$), pre-SFT LLMs' decision (P_1), and theoretical solution ($\hat{P}_1$) on P_3.

α_1 and θ_1 in $\hat{\mathcal{S}}_{\alpha_1} = \{0.05, 0.10, \ldots, 0.50\}$ and $\hat{\mathcal{S}}_{\theta_1} = \{1\times10^{-8}, 2\times10^{-8}, \ldots, 1\times10^{-7}\}$. Using the same method in Section 3.2, we set the investment attributes through two questions in natural language easy for LLMs to understand, rather than directly telling them the specific values of the parameters. For each investment attribute, we first calculate the theoretical optimal decision $\{\hat{P}_1(t)\}_{t\in\mathcal{T}}$ using (11) and Algorithm 1, and then calculate the investment proportion $\{\hat{P}_1(t)/X_1(t)\}_{t\in\mathcal{T}}$ using (1). Note that there exists a random perturbation $\{W(t)\}_{t\in\mathcal{T}}$ in (1), and we repeat 10 trials for each investment attribute. In summary, the SFT dataset contains 10^3 training samples.

4.2 Analysis of Parameter Convergence Rates

We theoretically show that fine-tuning LLMs on the training datasets constructed from theoretical solutions leads to faster parameter convergence compared to using real-user data. To ensure mathematical tractability, we make the following assumptions. First, when calculating the loss function, we only consider the values of the LLM's investment decision, theoretical solution, and real-user data. This is because the natural language parts for all three experiments are the same. Second, we assume that the sample size of the training dataset constructed from the theoretical solution and real-user data are both sufficiently large. Third, we assume that the output layer of the LLM is a Sigmoid layer, i.e., $\mathrm{Sigmoid}(\mathbf{z}) = [1+\exp(-\mathbf{w}^\top\mathbf{z})]^{-1}$, where $\mathbf{w}$ is the model parameter and $\mathbf{z}$ is the output layer's input. Although the output layer of the LLM may be more complex, this simplification makes the theoretical analysis tractable. We denote the ranges of the LLM's investment decision $P_1(t)$, theoretical solution $\hat{P}_1(t)$, and real-user data $\tilde{P}_1(t)$ as $\mathcal{P}_1(t)$, $\hat{\mathcal{P}}_1(t)$, and $\tilde{\mathcal{P}}_1(t)$, respectively.

Given the above assumptions, in the following, we analyze the parameter convergence rate in fine-tuning. First, according to the second assumption, when fine-tuning the LLM using the training dataset constructed from the theoretical solution, we can express the cross-entropy loss function as

$$\hat{L}(\mathbf{w}) = -\sum_{t\in\mathcal{T}} \int_{\hat{\mathcal{P}}_1(t)} f_{\hat{P}_1(t)}(x) \ln f_{P_1(t)}(x)\mathrm{d}x, \quad (6)$$

where $f_{P_1(t)}(\cdot)$ and $f_{\hat{P}_1(t)}(\cdot)$ represent the probability density functions of $P_1(t)$ and $\hat{P}_1(t)$ in the training dataset, respectively. Similarly, we can define the cross-entropy loss function $\tilde{L}(\mathbf{w})$ when fine-tuning the LLM using the real-user data.

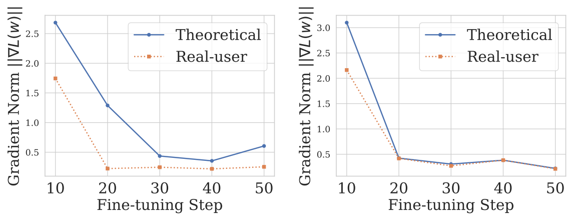
Next, we derive the analytical form of $f_{\hat{P}_1(t)}(\cdot)$ and $f_{\tilde{P}_1(t)}(\cdot)$. When we construct the SFT dataset, we uniformly set the values of the risk aversion coefficient α_1 and the influence coefficient θ_1 within a rectangular region. Therefore, we assume that α_1 and θ_1 satisfy two uniform distributions. As shown in in Appendix A.7, we can prove that

$$f_{\hat{P}_1(t)}(x) \approx cx^{-2},\ x\in\hat{\mathcal{P}}_1(t), \quad (7)$$

where c is a normalization parameter. Equation (7) is consistent with the empirical research in the field of behavioral finance, which shows that the distribution of trading volume often exhibits a power-law characteristic (Iori, 2002). From (7), we can find that the probability distribution function $f_{\hat{P}_1(t)}(\cdot)$ is a monotonically decreasing function. Thus, we assume that $f_{P_1(t)}(\cdot)$ is also monotonically decreasing. Because the real-user data often have a bigger noise than the theoretical solution, we assume that $\tilde{P}_1(t)$ is $\hat{P}_1(t)$ plus a white noise $n(t)$ that independently and identically satisfies a uniform distribution $U(-\varepsilon,\varepsilon)$, i.e, $\tilde{P}_1(t) = \hat{P}_1(t) + n(t)$. Then, as shown in Appendix A.8, we can derive the expression of $f_{\tilde{P}_1(t)}(\cdot)$. Finally, given $f_{\hat{P}_1(t)}(\cdot)$ and $f_{\tilde{P}_1(t)}(\cdot)$, we can calculate the gradient norms of the loss function and further prove that

$$\|\nabla\hat{L}(\mathbf{w})\| > \|\nabla\tilde{L}(\mathbf{w})\|. \quad (8)$$

That is, the gradient norm when using the training dataset constructed from the theoretical solution



(a) `Qwen-2` (b) `Llama-3.1`

Figure 3: Comparison of the gradient norms between using theoretical solution and real-user data.

is higher than when using real-user data. This is because, once the parameters are given, the real-user data are noisy, while the theoretical solution is deterministic. According to (Chen and Yang, 2018), the gradient descent algorithm converges faster when the gradient norm is larger. Thus, from (8), we can draw the conclusion that the gradient descent algorithm converges faster when using the training dataset compared to using real-user data.

We conduct an experiment to validate our above analysis on open-source models including `Qwen-2` and `Llama-3.1`. We construct the SFT datasets using both the theoretical solution and real-user data, and fine-tune the LLMs with these training datasets using low-rank adaptation (LoRA) in (Hu et al., 2021). We set the LoRA rank, alpha, and dropout rate as 4, 32, and 0.1, respectively, and keep the training parameters, such as the learning rate and batch size, etc., unchanged. The experimental results of the gradient norm $\|\nabla L(\mathbf{w})\|$ are in Figure 3. From Figure 3, the gradient norm when using the training dataset constructed from theoretical solution is significantly higher than when using real-user data across different LLMs, validating that fine-tuning LLMs on the training datasets constructed from theoretical solution leads to faster parameter convergence than using real-user data.

5 Experiments & Performance Validation

To answer Q_4, we conduct experiments to verify the alignment performance of **InvestAgent**s with real-user data in the simple problem P_3 and the original complex problems P_1 and P_2.

5.1 Performance of InvestAgent in P_3

Experimental Setup. To compare the alignment performance of pre-SFT LLMs and **InvestAgent**s with real-user data, we develop a Python-based simulation environment. The prompt we used is in Figure 11 in Appendix A.13. For different investment attributes, we select α_1 from $\mathcal{S}_{\alpha_1} = \{0.09, 0.13, 0.19, 0.26, 0.38\}$ and θ_1 from $\mathcal{S}_{\theta_1} = \{0, 1\times10^{-8}, \ldots, 1\times10^{-7}\}$. Given $\{W(t)\}_{t\in\mathcal{T}}$ in (1), we use 10 random seeds for each investment attribute, producing a total of 550 trials.

Experimental Results. Similarly to the data-processing method in Section 3.3, we plot the mean and the 95% confidence interval of the real-user data, denoted by $\tilde{P}_1$, and the pre-SFT LLMs' and **InvestAgent**s' investment decisions based on 10 repeated trials with the corresponding investment attribute, denoted by P_1, and P_1^{SFT}, respectively. We also plot the theoretical solutions, denoted by $\hat{P}_1$. The experimental results are in Figure 4. Here, we take the investment attribute $\alpha_1 = 0.13$ and $\theta_1 = 7\times10^{-8}$ as an example, and we observe the same trend for other values. As shown in Figure 4, **InvestAgent**s' investment decisions are significantly closer to real-user data and theoretical solutions than pre-SFT LLMs across different LLMs.

Additionally, to quantitatively evaluate how **InvestAlign** can help pre-SFT LLMs align with real-user data in P_3, we calculate the overall MSE between the mean of pre-SFT LLMs' decisions and real-user data, which is

$$\text{Overall MSE}(P_1, \tilde{P}_1) = \frac{1}{|\mathcal{M}||\mathcal{N}||\mathcal{T}|} \cdot \sum_{m\in\mathcal{M}, n\in\mathcal{N}, t\in\mathcal{T}} [P_1^{mn}(t) - \tilde{P}_1^{mn}(t)]^2, \tag{9}$$

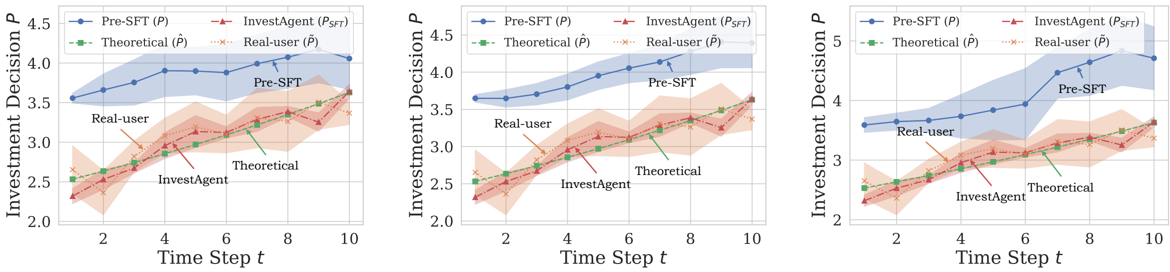
and the overall MSE between the mean of **InvestAgent**s' decisions and real-user data, which is

$$\text{Overall MSE}(P_1^{\mathrm{SFT}}, \tilde{P}_1) = \frac{1}{|\mathcal{M}||\mathcal{N}||\mathcal{T}|} \cdot \sum_{m\in\mathcal{M}, n\in\mathcal{N}, t\in\mathcal{T}} [P_1^{\mathrm{SFT},mn}(t) - \tilde{P}_1^{mn}(t)]^2, \tag{10}$$

where $\{\tilde{P}_{mn}(t)\}_{t\in\mathcal{T}}$ represents the mean of the real-user data in class $\mathcal{I}^{mn}$, $\{P_{mn}(t)\}_{t\in\mathcal{T}}$ and $\{P_{SFT,mn}(t)\}_{t\in\mathcal{T}}$ represents the mean of the pre-SFT LLMs' and **InvestAgent**s' investment decisions with the corresponding investment attribute, respectively. The experimental results are in Table 1. As shown in Table 1, **InvestAlign** helps reduce the overall MSEs by $45.59\% \sim 61.26\%$.

Furthermore, we also conduct an ablation study on the hyper-parameters of fine-tuning, including LoRA Rank and fine-tuning steps, as shown in Appendix A.9. We find that the overall MSE decreases as either LoRA Rank or fine-tuning steps increase.

The above results validate the effectiveness of our proposed method **InvestAlign**, i.e., fine-tuning LLMs using the SFT dataset constructed from the theoretical solution can align them better with investor decision-making under herd behavior.



(a) `GPT-3.5` (b) `Qwen-2` (c) `Llama-3.1`

Figure 4: Comparison of real-user data ($\tilde{P}_1$), pre-SFT LLMs' investment decision (P_1), **InvestAgent**s' investment decision (P_1^{SFT}), and theoretical solution ($\hat{P}_1$) on P_3.

Overall MSE		Pre-SFT	InvestAgent	Reduction
P_3	`GPT-3.5`	4.44	1.72	-61.26%
	`Qwen-2`	3.97	2.16	-45.59%
	`Llama-3.1`	4.08	1.59	-61.03%
P_1	`GPT-3.5`	14.03	7.46	-46.84%
	`Qwen-2`	17.22	7.46	-56.69%
	`Llama-3.1`	13.07	7.25	-44.52%
P_2	`Qwen-2`	15.66	6.12	-60.92%
	`Llama-3.1`	12.28	6.66	-45.77%

Table 1: Comparison of the overall MSE between pre-SFT LLMs' and **InvestAgent**s' investment decisions with real-user data in P_1, P_2, and P_3.

5.2 Performance of InvestAgent in P_1 and P_2

Experimental Setup. This experiment shows the alignment performance of our proposed **InvestAlign**, i.e., using LLMs fine-tuned from P_3 to solve P_1 and P_2. The prompts we use in P_1 and P_2 are in Figure 13 and Figure 14 in Appendix A.13, respectively. The investment attributes are set the same as those in Section 5.1.

Also, we collect 80 and 44 real-user data for P_1 and P_2, respectively, and the participants are also primarily professionals and students in the fields of finance to reduce bias and noise. The questionnaires we use in P_1 and P_2 are in Figure 16 and Figure 17 in Appendix A.14, respectively.

Experimental Results. Using the same method in Section 5.1, we list the overall MSE between the mean of pre-SFT LLMs' investment decisions with real-user data, Overall MSE$(P_1, \tilde{P}_1)$, and the overall MSE between the mean of **InvestAgent**s' investment decisions with real-user data, Overall MSE$(P_1^{\text{SFT}}, \tilde{P}_1)$, for P_1 in Table 1. For P_2, we list Overall MSE$(P_1 \cup P_2, \tilde{P}_1 \cup \tilde{P}_2)$ and Overall MSE$(P_1^{\text{SFT}} \cup P_2^{\text{SFT}}, \tilde{P}_1 \cup \tilde{P}_2)$ correspondingly. As shown in Table 1, **InvestAlign** helps reduce the overall MSEs by $44.53\% \sim 56.68\%$ and $45.77\% \sim 60.92\%$ in P_1 and P_2, respectively. The experimental results validate the effectiveness of our proposed **InvestAlign**, and show that the **InvestAgent**s fine-tuned using the theoretical solution in a similar and simpler problem can better align with human decision-making processes in a complex problem than pre-SFT LLMs. It demonstrates the potential of **InvestAlign** to solve complex optimal investment problems and align LLMs with investor decision-making processes.

In addition to the experiments mentioned above, we also: 1) supplement smaller samples of real-user data with theoretical solutions to construct a training dataset to improve robustness; 2) compare **InvestAgent**s with LLMs fine-tuned using the baseline FinGPT dataset (Yang et al., 2023a); and 3) validate that **InvestAgent**s can reflect economic principles in the presence of investor herd behavior. The experimental results and analyses are in Appendices A.10–A.12, respectively.

6 Conclusion

Studying investor decision-making processes under the herd behavior is of great significance in microeconomics and behavioral finance. LLMs can be leveraged to assist in solving complex investment problems. To fine-tune LLMs for alignment with human decision-making processes, a substantial amount of real-user data is required. However, the cost of collecting the real-user data is high, and there are concerns regarding privacy and security. To address data scarcity, we propose **InvestAlign**, a novel method that constructs training datasets using the theoretical solution of a similar and simple problem to align LLMs with investor behavior. We demonstrate that fine-tuning LLMs on these training datasets leads to faster parameter convergence compared to using real-user data. The experimental results indicate that **InvestAgent**s, fine-tuned with **InvestAlign**, achieves superior alignment performance in the original complex problem.

Limitations

As a preliminary work in this field, **InvestAlign** does not claim universal applicability for all complex optimal investment problems, and the theoretical solution may not fully capture all nuances of real-world investor behavior. We aim to address a specific challenge: data scarcity when training LLMs in the context of investor decision-making. In our future work, we plan to: 1) extend InvestAlign to diverse behavioral biases, e.g., overconfidence and loss aversion; 2) incorporate RLHF on **InvestAgent**s to complement SFT to assess the effectiveness of different alignment methods in investment decision-making tasks.

Ethics Statement

In our study, participants were primarily recruited from professional and student populations within the fields of microeconomics and behavioral finance. To ensure diversity and representativeness, we employed targeted recruitment strategies, such as reaching out through academic institutions and professional networks. Regarding compensation, participants were remunerated according to the standard rates for similar studies in the respective regions. Payment levels were carefully considered based on participants' demographic characteristics to ensure fair compensation for their time and expertise. Participants were fully informed about the study's objectives, how their data would be used, and their rights to withdraw from the study at any time without penalty. We have thoroughly reviewed the real-user data to ensure it is free from potential discrimination, human rights violations, bias, exploitation, or any other ethical concerns. Additionally, the data does not contain any information that could identify individuals or include offensive content. The data collection protocol is approved (or determined exempt) by an ethics review board.

References

Tyler Abbot. 2017. Heterogeneous risk preferences in financial markets. In *Paris December 2016 Finance Meeting EUROFIDAI-AFFI*.

Josh Achiam, Steven Adler, Sandhini Agarwal, Lama Ahmad, Ilge Akkaya, Florencia Leoni Aleman, Diogo Almeida, Janko Altenschmidt, Sam Altman, Shyamal Anadkat, et al. 2023. Gpt-4 technical report. *arXiv preprint arXiv:2303.08774*.

Maqsood Ahmad and Qiang Wu. 2022. Does herding behavior matter in investment management and perceived market efficiency? Evidence from an emerging market. *Management Decision*, 60(8):2148–2173.

Siyu An, Qin Li, Junru Lu, Di Yin, and Xing Sun. 2024. FinVerse: An autonomous agent system for versatile financial analysis. *arXiv preprint arXiv:2406.06379*.

Yuntao Bai, Andy Jones, Kamal Ndousse, Amanda Askell, Anna Chen, Nova DasSarma, Dawn Drain, Stanislav Fort, Deep Ganguli, Tom Henighan, et al. 2022. Training a helpful and harmless assistant with reinforcement learning from human feedback. *arXiv preprint arXiv:2204.05862*.

Yejin Bang, Samuel Cahyawijaya, Nayeon Lee, Wenliang Dai, Dan Su, Bryan Wilie, Holy Lovenia, Ziwei Ji, Tiezheng Yu, Willy Chung, et al. 2023. A multitask, multilingual, multimodal evaluation of ChatGPT on reasoning, hallucination, and interactivity. *arXiv preprint arXiv:2302.04023*.

Sushil Bikhchandani and Sunil Sharma. 2000. Herd behavior in financial markets. *IMF Staff papers*, 47(3):279–310.

Daniil A Boiko, Robert MacKnight, and Gabe Gomes. 2023. Emergent autonomous scientific research capabilities of large language models. *arXiv preprint arXiv:2304.05332*.

Samuel R Bowman, Jeeyoon Hyun, Ethan Perez, Edwin Chen, Craig Pettit, Scott Heiner, Kamilė Lukošiūtė, Amanda Askell, Andy Jones, Anna Chen, et al. 2022. Measuring progress on scalable oversight for large language models. *arXiv preprint arXiv:2211.03540*.

Jeffrey R Brown, Zoran Ivković, Paul A Smith, and Scott Weisbenner. 2008. Neighbors matter: Causal community effects and stock market participation. *The Journal of Finance*, 63(3):1509–1531.

Jianting Chen and Xiang Yang. 2018. Survey of unstable gradients in deep neural network training. *Journal of Software*, 29(7):2071–2091.

Yiting Chen, Tracy Xiao Liu, You Shan, and Songfa Zhong. 2023. The emergence of economic rationality of GPT. *Proceedings of the National Academy of Sciences*, 120(51):e2316205120.

Paul F Christiano, Jan Leike, Tom Brown, Miljan Martic, Shane Legg, and Dario Amodei. 2017. Deep reinforcement learning from human preferences. *Advances in neural information processing systems*, 30.

Ning Ding, Yulin Chen, Bokai Xu, Yujia Qin, Zhi Zheng, Shengding Hu, Zhiyuan Liu, Maosong Sun, and Bowen Zhou. 2023. Enhancing chat language models by scaling high-quality instructional conversations. *arXiv preprint arXiv:2305.14233*.

Abhimanyu Dubey, Abhinav Jauhri, Abhinav Pandey, Abhishek Kadian, Ahmad Al-Dahle, Aiesha Letman, Akhil Mathur, Alan Schelten, Amy Yang, Angela Fan, et al. 2024. The llama 3 herd of models. *arXiv preprint arXiv:2407.21783*.

Chen Gao, Xiaochong Lan, Nian Li, Yuan Yuan, Jingtao Ding, Zhilun Zhou, Fengli Xu, and Yong Li. 2024. Large language models empowered agent-based modeling and simulation: A survey and perspectives. *Humanities and Social Sciences Communications*, 11(1):1–24.

Wayne Geerling, G Dirk Mateer, Jadrian Wooten, and Nikhil Damodaran. 2023. ChatGPT has aced the test of understanding in college economics: Now what? *The American Economist*, 68(2):233–245.

Navid Ghaffarzadegan, Aritra Majumdar, Ross Williams, and Niyousha Hosseinichimeh. 2023. Generative agent-based modeling: Unveiling social system dynamics through coupling mechanistic models with generative artificial intelligence. *arXiv preprint arXiv:2309.11456*.

Team GLM, Aohan Zeng, Bin Xu, Bowen Wang, Chenhui Zhang, Da Yin, Diego Rojas, Guanyu Feng, Hanlin Zhao, Hanyu Lai, Hao Yu, Hongning Wang, Jiadai Sun, Jiajie Zhang, Jiale Cheng, Jiayi Gui, Jie Tang, Jing Zhang, Juanzi Li, Lei Zhao, Lindong Wu, Lucen Zhong, Mingdao Liu, Minlie Huang, Peng Zhang, Qinkai Zheng, Rui Lu, Shuaiqi Duan, Shudan Zhang, Shulin Cao, Shuxun Yang, Weng Lam Tam, Wenyi Zhao, Xiao Liu, Xiao Xia, Xiaohan Zhang, Xiaotao Gu, Xin Lv, Xinghan Liu, Xinyi Liu, Xinyue Yang, Xixuan Song, Xunkai Zhang, Yifan An, Yifan Xu, Yilin Niu, Yuantao Yang, Yueyan Li, Yushi Bai, Yuxiao Dong, Zehan Qi, Zhaoyu Wang, Zhen Yang, Zhengxiao Du, Zhenyu Hou, and Zihan Wang. 2024. Chatglm: A family of large language models from glm-130b to glm-4 all tools. *Preprint*, arXiv:2406.12793.

Tian Guo and Emmanuel Hauptmann. 2024. Fine-tuning large language models for stock return prediction using newsflow. *arXiv preprint arXiv:2407.18103*.

John J Horton. 2023. Large language models as simulated economic agents: What can we learn from homo silicus? Technical report, National Bureau of Economic Research.

Edward J Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. 2021. Lora: Low-rank adaptation of large language models. *arXiv preprint arXiv:2106.09685*.

Giulia Iori. 2002. A microsimulation of traders activity in the stock market: the role of heterogeneity, agents' interactions and trade frictions. *Journal of Economic Behavior & Organization*, 49(2):269–285.

Jiaming Ji, Tianyi Qiu, Boyuan Chen, Borong Zhang, Hantao Lou, Kaile Wang, Yawen Duan, Zhonghao He, Jiayi Zhou, Zhaowei Zhang, et al. 2023. Ai alignment: A comprehensive survey. *arXiv preprint arXiv:2310.19852*.

Grgur Kovač, Rémy Portelas, Peter Ford Dominey, and Pierre-Yves Oudeyer. 2023. The socialai school: Insights from developmental psychology towards artificial socio-cultural agents. *arXiv preprint arXiv:2307.07871*.

Jean Lee, Nicholas Stevens, Soyeon Caren Han, and Minseok Song. 2024. A survey of large language models in finance (finLLMs). *arXiv preprint arXiv:2402.02315*.

Jiaxiang Li, Siliang Zeng, Hoi-To Wai, Chenliang Li, Alfredo Garcia, and Mingyi Hong. 2024a. Getting more juice out of the SFT data: Reward learning from human demonstration improves SFT for LLM alignment. *arXiv preprint arXiv:2405.17888*.

Lezhi Li, Ting-Yu Chang, and Hai Wang. 2023. Multimodal Gen-AI for fundamental investment research. *arXiv preprint arXiv:2401.06164*.

Nian Li, Chen Gao, Mingyu Li, Yong Li, and Qingmin Liao. 2024b. Econagent: Large language model-empowered agents for simulating macroeconomic activities. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 15523–15536.

Xuan Liu, Jie Zhang, Song Guo, Haoyang Shang, Chengxu Yang, and Quanyan Zhu. 2024. Exploring prosocial irrationality for LLM agents: A social cognition view. *arXiv preprint arXiv:2405.14744*.

Andres M. Bran, Sam Cox, Oliver Schilter, Carlo Baldassari, Andrew D White, and Philippe Schwaller. 2024. Augmenting large language models with chemistry tools. *Nature Machine Intelligence*, pages 1–11.

Robert C Merton. 1969. Lifetime portfolio selection under uncertainty: The continuous-time case. *The Review of Economics and Statistics*, pages 247–257.

Ethan Perez, Sam Ringer, Kamilė Lukošiūtė, Karina Nguyen, Edwin Chen, Scott Heiner, Craig Pettit, Catherine Olsson, Sandipan Kundu, Saurav Kadavath, et al. 2022. Discovering language model behaviors with model-written evaluations. *arXiv preprint arXiv:2212.09251*.

John W Pratt. 1978. Risk aversion in the small and in the large. In *Uncertainty in economics*, pages 59–79. Elsevier.

Alfréd Rényi. 2007. *Probability theory*. Courier Corporation.

Leonard CG Rogers. 2013. *Optimal investment*, volume 1007. Springer.

Jun Shao. 2008. *Mathematical statistics*. Springer Science & Business Media.

Jacob Steinhardt. 2023. Emergent deception and emergent optimization. *Bounded Regret*, 19:2023.

Cheng Wang, Chuwen Wang, Yu Zhao, Shirong Zeng, Wang Zhang, and Ronghui Ning. 2024a. Behavioral simulation: Exploring a possible next paradigm for science. *arXiv preprint arXiv:2401.09851*.

Huisheng Wang and H Vicky Zhao. 2024a. Optimal investment under the influence of decision-changing imitation. In *2024 China Automation Congress (CAC)*, pages 8252–8257. IEEE.

Huisheng Wang and H Vicky Zhao. 2024b. Optimal investment with herd behaviour using rational decision decomposition. In *2024 43rd Chinese Control Conference (CCC)*, pages 1645–1651. IEEE.

Huisheng Wang and H Vicky Zhao. 2025. Optimal investment under mutual strategy influence among agents. *arXiv preprint arXiv:2501.14259*.

Junqi Wang, Chunhui Zhang, Jiapeng Li, Yuxi Ma, Lixing Niu, Jiaheng Han, Yujia Peng, Yixin Zhu, and Lifeng Fan. 2024b. Evaluating and modeling social intelligence: A comparative study of human and AI capabilities. *arXiv preprint arXiv:2405.11841*.

Saizhuo Wang, Hang Yuan, Leon Zhou, Lionel M Ni, Heung-Yeung Shum, and Jian Guo. 2023a. Alphagpt: Human-ai interactive alpha mining for quantitative investment. *arXiv preprint arXiv:2308.00016*.

Yizhong Wang, Hamish Ivison, Pradeep Dasigi, Jack Hessel, Tushar Khot, Khyathi Chandu, David Wadden, Kelsey MacMillan, Noah A Smith, Iz Beltagy, et al. 2023b. How far can camels go? Exploring the state of instruction tuning on open resources. *Advances in Neural Information Processing Systems*, 36:74764–74786.

Zhichao Wang, Bin Bi, Can Huang, Shiva Kumar Pentyala, Zixu James Zhu, Sitaram Asur, and Na Claire Cheng. 2024c. UNA: Unifying alignments of RLHF/PPO, DPO and KTO by a generalized implicit reward function. *arXiv preprint arXiv:2408.15339*.

Zhichao Wang, Bin Bi, Shiva Kumar Pentyala, Kiran Ramnath, Sougata Chaudhuri, Shubham Mehrotra, Xiang-Bo Mao, Sitaram Asur, et al. 2024d. A comprehensive survey of LLM alignment techniques: RLHF, RLAIF, PPO, DPO and more. *arXiv preprint arXiv:2407.16216*.

Shijie Wu, Ozan Irsoy, Steven Lu, Vadim Dabravolski, Mark Dredze, Sebastian Gehrmann, Prabhanjan Kambadur, David Rosenberg, and Gideon Mann. 2023. Bloomberggpt: A large language model for finance. *arXiv preprint arXiv:2303.17564*.

Yuhang Xiao, Yudi Lin, and Ming-Chang Chiu. 2024. Behavioral bias of vision-language models: A behavioral finance view. In *ICML 2024 Workshop on LLMs and Cognition*.

Shiming Xie, Hong Chen, Fred Yu, Zeye Sun, and Xiuyu Wu. 2024. Minor SFT loss for LLM fine-tune to increase performance and reduce model deviation. *arXiv preprint arXiv:2408.10642*.

An Yang, Baosong Yang, Binyuan Hui, Bo Zheng, Bowen Yu, Chang Zhou, Chengpeng Li, Chengyuan Li, Dayiheng Liu, Fei Huang, Guanting Dong, Haoran Wei, Huan Lin, Jialong Tang, Jialin Wang, Jian Yang, Jianhong Tu, Jianwei Zhang, Jianxin Ma, Jin Xu, Jingren Zhou, Jinze Bai, Jinzheng He, Junyang Lin, Kai Dang, Keming Lu, Keqin Chen, Kexin Yang, Mei Li, Mingfeng Xue, Na Ni, Pei Zhang, Peng Wang, Ru Peng, Rui Men, Ruize Gao, Runji Lin, Shijie Wang, Shuai Bai, Sinan Tan, Tianhang Zhu, Tianhao Li, Tianyu Liu, Wenbin Ge, Xiaodong Deng, Xiaohuan Zhou, Xingzhang Ren, Xinyu Zhang, Xipin Wei, Xuancheng Ren, Yang Fan, Yang Yao, Yichang Zhang, Yu Wan, Yunfei Chu, Yuqiong Liu, Zeyu Cui, Zhenru Zhang, and Zhihao Fan. 2024. Qwen2 technical report. *arXiv preprint arXiv:2407.10671*.

Hongyang Yang, Xiao-Yang Liu, and Christina Dan Wang. 2023a. Fingpt: Open-source financial large language models. *arXiv preprint arXiv:2306.06031*.

Yi Yang, Yixuan Tang, and Kar Yan Tam. 2023b. Investlm: A large language model for investment using financial domain instruction tuning. *arXiv preprint arXiv:2309.13064*.

Shengyu Zhang, Linfeng Dong, Xiaoya Li, Sen Zhang, Xiaofei Sun, Shuhe Wang, Jiwei Li, Runyi Hu, Tianwei Zhang, Fei Wu, et al. 2023. Instruction tuning for large language models: A survey. *arXiv preprint arXiv:2308.10792*.

Xuanyu Zhang and Qing Yang. 2023. Xuanyuan 2.0: A large chinese financial chat model with hundreds of billions parameters. In *Proceedings of the 32nd ACM international conference on information and knowledge management*, pages 4435–4439.

Shujuan Zhao, Lingfeng Qiao, Kangyang Luo, Qian-Wen Zhang, Junru Lu, and Di Yin. 2024. SNFinLLM: Systematic and nuanced financial domain adaptation of Chinese large language models. *arXiv preprint arXiv:2408.02302*.

Wayne Xin Zhao, Kun Zhou, Junyi Li, Tianyi Tang, Xiaolei Wang, Yupeng Hou, Yingqian Min, Beichen Zhang, Junjie Zhang, Zican Dong, et al. 2023. A survey of large language models. *arXiv preprint arXiv:2303.18223*.

Shifen Zhou and Xiaojun Liu. 2022. Internet postings and investor herd behavior: Evidence from China's open-end fund market. *Humanities and Social Sciences Communications*, 9(1):1–11.

A Appendix

A.1 Theoretical Optimal Investment Decisions of P_3

A_1's optimal investment decision for P_3 is given by

$$\hat{P}_1(t) = \frac{\alpha_2\sigma^2\eta\exp[2r(T-t)]+\theta_1}{\alpha_1\sigma^2\eta\exp[2r(T-t)]+\theta_1}\cdot\frac{v}{\alpha_2\sigma^2}\exp[r(t-T)],\ t\in\mathcal{T}, \tag{11}$$

where the parameter η can be numerically calculated using Algorithm 1, and A_2's optimal investment decision for P_3 is given by

$$\hat{P}_2(t) = \frac{v}{\alpha_2\sigma^2}\exp[r(t-T)],\ t\in\mathcal{T}. \tag{12}$$

The proof can be found in (Wang and Zhao, 2024b).

Algorithm 1: Numerical Method of the Parameter η in P_3.

Input: Interest rate: r;
Excess return rate: v;
Volatility: σ;
Initial fund: $x_{1,0}$;
Risk aversion coefficients: α_1 and α_2;
Investment period: T;
Influence coefficient: θ_1;
Error tolerance: ε.
Output: The parameter η.

$\eta_0 = \exp\left[-\alpha_1 x_{1,0}\mathrm{e}^{rT} - \frac{v^2T}{2\sigma^2}\right]$, $\Delta\eta_0 = +\infty$, $k=0$, $\vartheta = \frac{\theta}{\alpha_1\sigma^2}$;
while $\Delta\eta_k \geqslant \varepsilon$ **do**
  $\eta_{k+1} = \eta_0\exp\int_{\mathcal{T}}\frac{\vartheta^2v^2(\alpha_1/\alpha_2-1)^2\mathrm{d}t}{2\sigma^2\{\eta_k\exp[2r(T-t)]+\vartheta\}^2}$;
  $\Delta\eta_{k+1} = |\eta_{k+1}-\eta_k|$;
  $k \leftarrow k+1$;
end
$\eta \approx \eta_k$.

A.2 Parameter Settings

Following the work in (Wang and Zhao, 2024b), we set the default parameter values in this work as:
- The interest rate of the deposit: $r = 0.04$.
- The excess return rate of the stock: $v = 0.03$.
- The volatility of the stock: $\sigma = 0.17$.
- The investment period: $T = 10$.
- A_1's initial fund: $x_{1,0} = 10$.

Additionally, for P_1 and P_3, we further set:
- A_2's risk aversion coefficient: $\alpha_2 = 0.2$.

A.3 Calculation Methods of Investment Attributes

From the work in (Pratt, 1978), the risk aversion coefficient α_1 reflects the participant's preference between risky and risk-free options. If the participant is indifferent between the following two options: (1) receiving w_1 with probability p_i, and receiving nothing with probability $1-p_i$, or (2) receiving w_2, his/her risk aversion coefficient α_1^i can be determined by

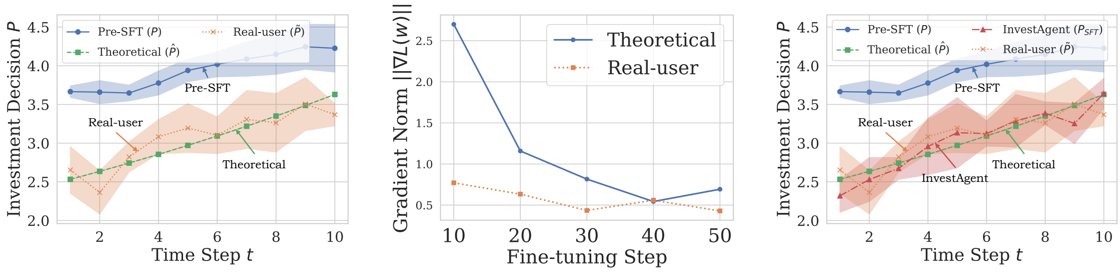
$$p_i = \frac{\exp(-\alpha_1^i w_2)-1}{\exp(-\alpha_1^i w_1)-1}. \tag{13}$$

We ask the participant to provide his/her response for p_i, from which we calculate his/her risk aversion coefficient α_1^i.

From the work in (Wang and Zhao, 2024b), the influence coefficient θ_1 quantifies the level of herd behavior. In the third part of the questionnaire, we ask participants: "On a scale $[0, 10]$, how much do you rely on the investment assistant when making decisions, where 10 represents the highest reliance level and 0 the lowest?" From the work in (Wang and Zhao, 2024b), the influence coefficient θ_1^i typically falls within the range $[0, 1 \times 10^{-7}]$. Thus, we set the participant's influence coefficient as $\theta_1^i = k_i \times 10^{-8}$, where k_i is the response.

A.4 Experimental Results of `GLM-4-9B-CHAT`

In this work, we also conduct experiments on `GLM-4-9B-CHAT` (GLM et al., 2024) for better comparison. The experimental results of `GLM-4-9B-CHAT` corresponding to Figures 2–4 are in Figure 5.

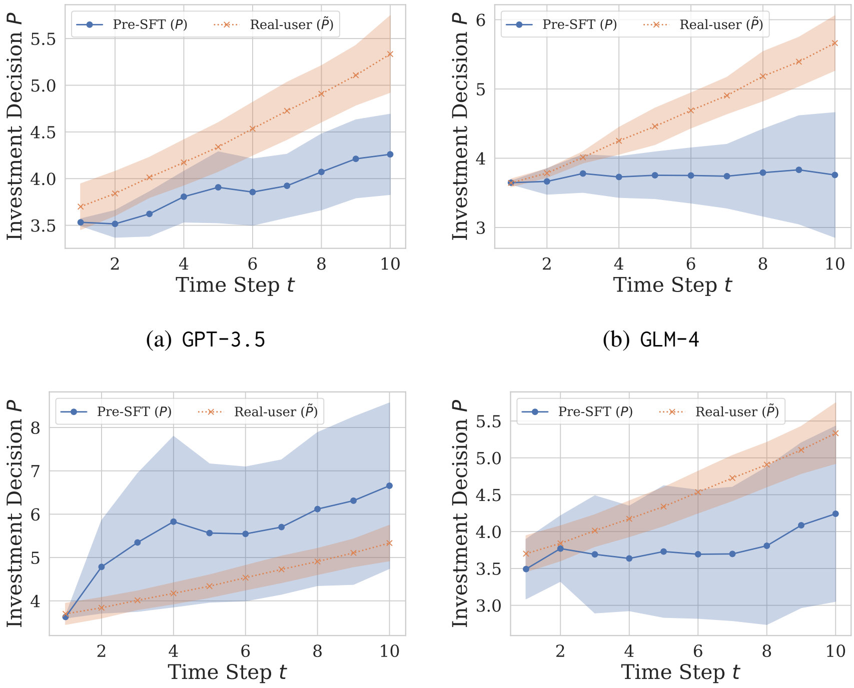


(a) Figure 2 (b) Figure 3 (c) Figure 4

Figure 5: The experimental results of `GLM-4-9B-CHAT` corresponding to Figures 2–4.

A.5 Comparison of Real-User Data and Pre-SFT LLMs' Investment Decision on P_1 and P_2

From Figure 6, we can find that the performances of pre-SFT LLMs in P_1 are misaligned with real-user data. Note that when A_2's influence coefficient $\theta_2 = 0$, the complex problem P_2 degenerates into the simpler problem P_3. Therefore, from Figure 2, we can also draw the conclusion that the performances of pre-SFT LLMs in P_2 are misaligned with real-user data.



(a) `GPT-3.5`

(b) `GLM-4`

(c) `Qwen-2`

(d) `Llama-3.1`

Figure 6: Comparison of real-user data ($\tilde{P}_1$) and pre-SFT LLMs' investment decision (P_1) on P_1.

A.6 Statistical Tests' Results of the Differences and Correlation Coefficients

We define the difference d_i and correlation coefficient ρ_i between $\{\tilde{P}_1^i(t)\}_{t\in\mathcal{T}}$ and $\{\hat{P}_1^i(t)\}_{t\in\mathcal{T}}$ as

$$d_i = \sum_{t\in\mathcal{T}} [\tilde{P}_1^i(t) - \hat{P}_1^i(t)], \tag{14}$$

and

$$\rho_i = \frac{\sum_{t\in\mathcal{T}} [\tilde{P}_1^i(t) - \bar{\tilde{P}}_1^i][\hat{P}_1^i(t) - \bar{\hat{P}}_1^i]}{\sqrt{\sum_{t\in\mathcal{T}} [\tilde{P}_1^i(t) - \bar{\tilde{P}}_1^i]^2 \sum_{t\in\mathcal{T}} [\hat{P}_1^i(t) - \bar{\hat{P}}_1^i]^2}}, \tag{15}$$

where $\bar{\tilde{P}}_1^i = \frac{1}{T}\sum_{t\in\mathcal{T}} \tilde{P}_1^i(t)$ and $\bar{\hat{P}}_1^i = \frac{1}{T}\sum_{t\in\mathcal{T}} \hat{P}_1^i(t)$, respectively.

For the differences $\{d_i\}_{i\in\mathcal{I}}$, the results show that their mean does not significantly deviate from 0 at the 1% significance level, with a t-statistic $= -1.075$. For the correlation coefficients $\{\rho_i\}_{i\in\mathcal{I}}$, the results show that their mean does not significantly deviate from 0.85 at the 1% significance level, with a t-statistic $= -0.843$. Since a mean difference close to 0 indicates minimal discrepancy and a correlation coefficient close to 0.85 reflects a strong positive relationship, we show that there exists significant consistency between the theoretical solution and real-user data.

A.7 Probability Distribution Function of A_1's Optimal Investment Decision of P_3

We assume the parameters α_1 and θ_1 satisfy two uniform distributions, denoted by $U(\underline{\alpha}, \overline{\alpha})$ and $U(\underline{\theta}, \overline{\theta})$, respectively. Therefore, their probability distribution functions are

$$f_{\alpha_1}(x) = \frac{1}{\overline{\alpha} - \underline{\alpha}},\ x \in [\underline{\alpha}, \overline{\alpha}],\ \text{and}\ f_{\theta_1}(x) = \frac{1}{\overline{\theta} - \underline{\theta}},\ x \in [\underline{\theta}, \overline{\theta}]. \tag{16}$$

From (11), using the convolution formula (Rényi, 2007), we have

$$
\begin{aligned}
f_{\hat{P}_1(t)}(x) = \frac{1}{\overline{\theta} - \underline{\theta}} \int_{\underline{\theta}}^{\overline{\theta}} f_{\alpha_1} \Bigg(& \frac{1}{\sigma^2 \eta \exp[2r(T-t)]} \\
& \cdot \left\{ \frac{\alpha_2 \sigma^2 \eta \exp[2r(T-t)] + y}{x} \cdot \frac{v}{\alpha_2 \sigma^2} \exp[r(t-T)] - y \right\} \Bigg) \\
& \cdot \frac{\alpha_2 \sigma^2 \eta \exp[2r(T-t)] + y}{\sigma^2 \eta \exp[2r(T-t)] x^2} \cdot \frac{v}{\alpha_2 \sigma^2} \exp[r(t-T)] \mathrm{d}y.
\end{aligned} \tag{17}
$$

Here, following the work in (Wang and Zhao, 2024a), we assume that η approximately remains constant when α_1 and θ_1 change slightly. Because $\hat{P}_1(t) \in \hat{\mathcal{P}}_1(t)$, we can rewrite (17) as

$$
f_{\hat{P}_1(t)}(x) \approx \frac{\min[\hat{\mathcal{P}}_1(t)] \cdot \max[\hat{\mathcal{P}}_1(t)]}{\max[\hat{\mathcal{P}}_1(t)] - \min[\hat{\mathcal{P}}_1(t)]} \cdot \frac{1}{x^2}, \ x \in \hat{\mathcal{P}}_1(t). \tag{18}
$$

That is, A_1's theoretical optimal decision $\hat{P}_1(t)$ approximately satisfies a Pareto distribution. To simplify the notation, we denote the normalization parameter as

$$
c = \frac{\min[\hat{\mathcal{P}}_1(t)] \cdot \max[\hat{\mathcal{P}}_1(t)]}{\max[\hat{\mathcal{P}}_1(t)] - \min[\hat{\mathcal{P}}_1(t)]}. \tag{19}
$$

A.8 Gradient Norms of the Loss Function

First, we derive the expression of $f_{\tilde{P}_1(t)}(\cdot)$ from (7). Using the convolution formula in (Rényi, 2007), we can obtain

$$
\begin{aligned}
f_{\tilde{P}_1(t)}(x) &\approx \frac{1}{2\varepsilon} \int_{-\varepsilon}^{\varepsilon} f_{\hat{P}_1(t)}(x-y) \mathrm{d}y \\
&= \begin{cases}
\frac{c}{2\varepsilon} \int_{\min[\hat{\mathcal{P}}_1(t)]-x}^{\varepsilon} \frac{1}{(x-y)^2} \mathrm{d}y, & x \in [\min[\hat{\mathcal{P}}_1(t)] - \varepsilon, \min[\hat{\mathcal{P}}_1(t)] + \varepsilon) \\
\frac{c}{2\varepsilon} \int_{-\varepsilon}^{\varepsilon} \frac{1}{(x-y)^2} \mathrm{d}y, & x \in [\min[\hat{\mathcal{P}}_1(t)] + \varepsilon, \max[\hat{\mathcal{P}}_1(t)] - \varepsilon) \\
\frac{c}{2\varepsilon} \int_{-\varepsilon}^{\max[\hat{\mathcal{P}}_1(t)]-x} \frac{1}{(x-y)^2} \mathrm{d}y, & x \in [\max[\hat{\mathcal{P}}_1(t)] - \varepsilon, \max[\hat{\mathcal{P}}_1(t)] + \varepsilon]
\end{cases} \\
&= \frac{c}{2\varepsilon} \left(\frac{1}{\max\{\min[\hat{\mathcal{P}}_1(t)], x - \varepsilon\}} - \frac{1}{\min\{\max[\hat{\mathcal{P}}_1(t)], x + \varepsilon\}} \right), \ x \in \tilde{\mathcal{P}}(t).
\end{aligned} \tag{20}
$$

Next, we calculate the gradient norms. We have

$$
\begin{aligned}
\nabla \hat{L}(\mathbf{w}) &= -\sum_{t \in \mathcal{T}} \int_{\hat{\mathcal{P}}_1(t)} f_{\hat{P}_1(t)}(x) \nabla \ln f_{P_1(t)}(x) \mathrm{d}x \\
&= -\sum_{t \in \mathcal{T}} \int_{\hat{\mathcal{P}}_1(t)} \frac{f_{\hat{P}_1(t)}(x)}{f_{P_1(t)}(x)} \nabla \mathrm{Sigmoid}(\mathbf{z}) \mathrm{d}x \\
&= -\mathbf{z} \sum_{t \in \mathcal{T}} \int_{\hat{\mathcal{P}}_1(t)} f_{\hat{P}_1(t)}(x) [1 - f_{P_1(t)}(x)] \mathrm{d}x \\
&= -\mathbf{z} \sum_{t \in \mathcal{T}} \left[\int_{\hat{\mathcal{P}}_1(t)} f_{\hat{P}_1(t)}(x) \mathrm{d}x - \int_{\hat{\mathcal{P}}_1(t)} f_{\hat{P}_1(t)}(x) f_{P_1(t)}(x) \mathrm{d}x \right] \\
&= -\mathbf{z} \sum_{t \in \mathcal{T}} \left[1 - \int_{\hat{\mathcal{P}}_1(t)} f_{\hat{P}_1(t)}(x) f_{P_1(t)}(x) \mathrm{d}x \right].
\end{aligned} \tag{21}
$$

Therefore, the gradient norm is

$$\|\nabla \hat{L}(\mathbf{w})\| = \|\mathbf{z}\| \sum_{t \in \mathcal{T}} \left[1 - \int_{\hat{\mathcal{P}}_1(t)} f_{\hat{P}_1(t)}(x) f_{P_1(t)}(x) \mathrm{d}x \right]. \tag{22}$$

Using the same method as above, we can obtain

$$\|\nabla \tilde{L}(\mathbf{w})\| = \|\mathbf{z}\| \sum_{t \in \mathcal{T}} \left[1 - \int_{\tilde{\mathcal{P}}_1(t)} f_{\tilde{P}_1(t)}(x) f_{P_1(t)}(x) \mathrm{d}x \right]. \tag{23}$$

Then, we compare the two gradient norms $\|\nabla \hat{L}(\mathbf{w})\|$ and $\|\nabla \tilde{L}(\mathbf{w})\|$. From (22) and (23), we only need to compare the following two integrals:

$$\hat{I} = \int_{\hat{\mathcal{P}}_1(t)} f_{\hat{P}_1(t)}(x) f_{P_1(t)}(x) \mathrm{d}x, \tag{24}$$

and

$$\tilde{I} = \int_{\tilde{\mathcal{P}}_1(t)} f_{\tilde{P}_1(t)}(x) f_{P_1(t)}(x) \mathrm{d}x. \tag{25}$$

Because the investment decisions of the pre-SFT LLM can be arbitrary due to the randomness of model parameters, we have

$$\hat{\mathcal{P}}_1(t) \subset \tilde{\mathcal{P}}_1(t) \subset \mathcal{P}_1(t). \tag{26}$$

Because $f_{P_1(t)}(\cdot)$ is monotonically decreasing, from (7) and (21), we can prove that

$$\hat{I} < \tilde{I} < 1, \tag{27}$$

and thus we have

$$\|\nabla \hat{L}(\mathbf{w})\| > \|\nabla \tilde{L}(\mathbf{w})\|. \tag{28}$$

A.9 The Ablation Study on the Hyper-parameters of SFT

We conduct an ablation study on the hyper-parameters of fine-tuning, including LoRA Rank and fine-tuning steps. Here, we take `Qwen-2` and `Llama-3.1` as examples. The experimental results are in Tables 2 and Table 3. It can be seen that our **InvestAlign** consistently enhances the agreement between the **InvestAgent** and real-user data across various hyperparameters. Furthermore, the overall MSE decreases as the strength of fine-tuning increases, either through a larger LoRA Rank or more fine-tuning steps, underscoring the effectiveness of **InvestAlign**. We hypothesize that full-parameter fine-tuning could yield even better results if computational resources permit, which we plan to explore in future studies.

Overall MSE	`Qwen-2`		
	$R = 4$	$R = 8$	$R = 16$
Pre-SFT	3.97	3.97	3.97
InvestAgent	3.09	2.16	1.35
Reduction	-22.17%	-45.60%	-65.99%
Overall MSE	`Llama-3.1`		
	$R = 4$	$R = 8$	$R = 16$
Pre-SFT	4.08	4.08	4.08
InvestAgent	2.40	1.59	1.36
Reduction	-41.18%	-61.03%	-66.67%

Table 2: Ablation study on the LoRA rank (R) using `Qwen-2` and `Llama-3.1`.

Overall MSE	Qwen-2				
	$S = 50$	$S = 100$	$S = 150$	$S = 200$	$S = 250$
Pre-SFT	3.97	3.97	3.97	3.97	3.97
InvestAgent	2.83	3.01	2.67	2.43	2.16
Reduction	-28.71%	-24.18%	-32.75%	-38.79%	-45.59%

Overall MSE	Llama-3.1				
	$S = 50$	$S = 100$	$S = 150$	$S = 200$	$S = 250$
Pre-SFT	4.08	4.08	4.08	4.08	4.08
InvestAgent	3.17	2.72	2.92	1.97	1.59
Reduction	-22.30%	-33.33%	-28.43%	-51.72%	-61.03%

Table 3: Ablation study on the fine-tuning step (S) using Qwen-2 and Llama-3.1.

A.10 The Experimental Results of Supplementing Smaller Samples of Real-user Data with Theoretical Solutions

Here, we take the complex problem P_1 and the simpler problem P_3 as examples. We conduct the experiments using the dataset of theoretical data and smaller samples of real-user data. The experimental results of P_1 and P_3 are in Table 4.

Overall MSE	Qwen-2	Llama-3.1
P_3: Absolute herd behavior with unilateral influence (simple problem)		
Pre-SFT LLM	3.97	4.08
Mix-SFT LLM (1:10)	2.85	3.17
Mix-SFT LLM (1:1)	2.38	1.76
Mix-SFT LLM (10:1)	**2.03**	1.64
InvestAgent	2.16	**1.59**
P_1: Relative herd behavior with unilateral influence (original complex problem)		
Pre-SFT LLM	17.22	13.07
Mix-SFT LLM (1:10)	11.33	10.68
Mix-SFT LLM (1:1)	9.65	8.98
Mix-SFT LLM (10:1)	**7.32**	**7.06**
InvestAgent	7.46	7.25

Table 4: Comparison of the overall MSE between pre-SFT LLMs', mix-SFT LLMs', and **InvestAgent**s' investment decisions with real-user data. "Mix-SFT LLM ($m : n$)" means that LLM was fine-tuned on a training dataset where the ratio of theoretical data to real-user data is m/n.

From Table 4, it can be observed that supplementing a portion of real-user data slightly improved the model's performance on average, i.e., **InvestAgent**s align more with real-user data, indicating that this approach can enhance the model's robustness to some extent. Notably, as the proportion of real-user data in the entire SFT dataset gradually increases, the robustness may improve, but the parameter convergence rate decreases. We have provided both theoretical and experimental evidence for this in Section 4.2.

A.11 The Experimental Results of Compare InvestAgents with LLMs Fine-tuned Using the Baseline FinGPT Dataset

Here, we take the complex problem P_1 and the simpler problem P_3 as examples. We conduct the experiments using the FinGPT datasets (Yang et al., 2023a), including `FinGPT-FinEval` and `FinGPT-ConvFinQA`, to fine-tune LLMs, and compare them with our proposed **InvestAgent**s. The experimental results of P_1, P_2, and P_3 are in Table 5.

Overall MSE	`Qwen-2`	`Llama-3.1`
P_3: Absolute herd behavior with unilateral influence (simple problem)		
Pre-SFT LLM	3.97	4.08
FinEval-SFT LLM	3.35	3.28
ConvFinQA-SFT LLM	2.77	1.96
InvestAgent	**2.16**	**1.59**
P_1: Relative herd behavior with unilateral influence (original complex problem)		
Pre-SFT LLM	17.22	13.07
FinEval-SFT LLM	13.74	11.16
ConvFinQA-SFT LLM	10.86	9.61
InvestAgent	**7.46**	**7.25**

Table 5: Comparison of the overall MSE between pre-SFT LLMs', FinEval-SFT LLMs', ConvFinQA-SFT LLMs' and **InvestAgent**s' investment decisions with real-user data.

From Table 5, it can be seen that **InvestAgent**s outperform the LLMs fine-tuned on the FinGPT datasets. This is because **InvestAgent**'s training dataset is specifically constructed for studying optimal investment problems with herding behavior, whereas FinGPT is more general. Therefore, **InvestAgent** shows better performance in the context of optimal investment analysis.

A.12 The Experimental Results of Validating that InvestAgents Better Reflect Economic Principles in the Presence of Investor Herd Behavior Compared to Pre-SFT LLMs

A.12.1 Experimental Setup

We conduct experiments to validate whether **InvestAgent**s better reflect established economic principles in scenarios involving investor herd behavior compared to pre-SFT LLMs. Our analysis focuses on two foundational economic hypotheses:

- H_1: In both unilateral and mutual influence scenarios, as the influence coefficient increases, which amplifies herd behavior, agents' investment decisions, i.e., $\{P_1(t)\}_{t\in\mathcal{T}}$ and $\{P_2(t)\}_{t\in\mathcal{T}}$, should exhibit progressive convergence (Wang and Zhao, 2024a, 2025, 2024b).

- H_2: In the mutual influence scenario, as the influence coefficient increases, which amplifies herd behavior, the mean of the sum of the two agents' terminal funds, i.e., $\mathbb{E}[X_1(T) + X_2(T)]$, should decrease (Wang and Zhao, 2025).

We choose A_1's influence coefficient θ_1 from $\hat{\mathcal{S}}_{\theta_1}$ or $\tilde{\mathcal{S}}_{\theta_1}$ for P_1 and P_3, and set the homogeneous influence coefficients of both A_1 and A_2, $\theta_1 = \theta_2 := \theta$, from $\hat{\mathcal{S}}_{\theta_1}$ or $\tilde{\mathcal{S}}_{\theta_1}$ for P_2. Other parameters are specified in Appendix A.2. We generate corresponding prompts and collect **InvestAgent**s' responses, from which we extract A_1's investment decision $\{P_1^{\mathrm{SFT}}(t)\}_{t\in\mathcal{T}}$ for P_1 and P_3, and both A_1's investment decision $\{P_1^{\mathrm{SFT}}(t)\}_{t\in\mathcal{T}}$ and A_2's investment decision $\{P_2^{\mathrm{SFT}}(t)\}_{t\in\mathcal{T}}$ for P_2, respectively.

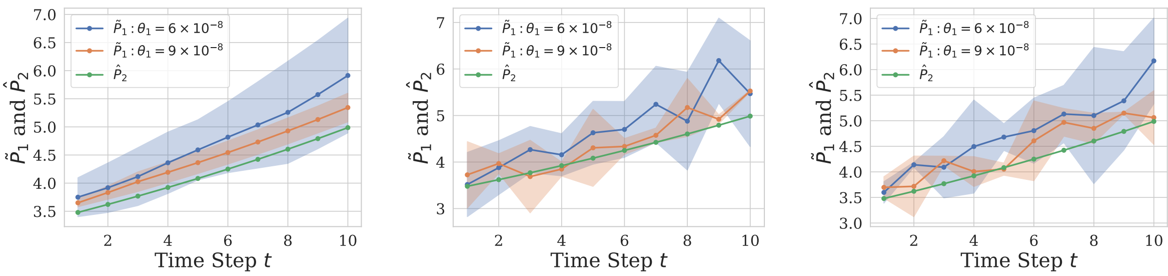
A.12.2 Validation of H_1

For P_1 and P_3, we plot A_1's decisions $\{P_1^{\mathrm{SFT}}(t)\}_{t\in\mathcal{T}}$ under different influence coefficients θ_1. To facilitate comparison of the herd behavior's impact on investment decisions, we simultaneously plot A_2's optimal decision $\{\hat{P}_2(t)\}_{t\in\mathcal{T}}$ calculated from (12). Additionally, we also show the results of real-user data for comparison. The experimental results are in Figure 7 and Figure 8, respectively. From Figure 7 and

Figure 8, we observe that as the influence coefficient θ_1 increases, A_1's decision $\{P_1^{\mathrm{SFT}}(t)\}_{t\in\mathcal{T}}$ gradually converges toward A_2's optimal decision $\{\hat{P}_2(t)\}_{t\in\mathcal{T}}$.

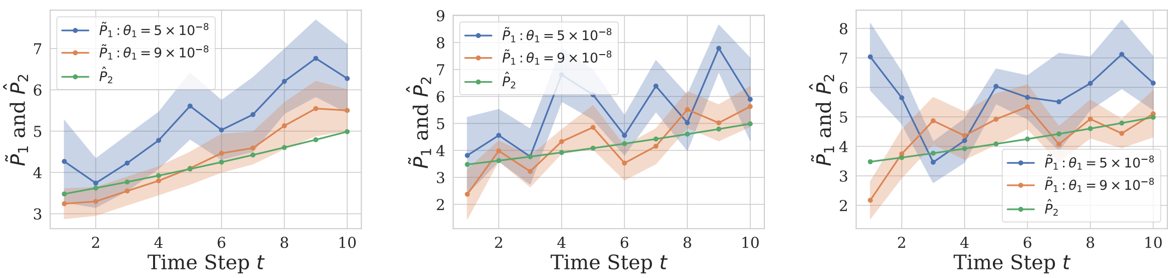
For P_2, we plot both A_1's decisions $\{P_1^{\mathrm{SFT}}(t)\}_{t\in\mathcal{T}}$ and A_2's decisions $\{P_2^{\mathrm{SFT}}(t)\}_{t\in\mathcal{T}}$ under different influence coefficients θ in Figure 9. From Figure 9, we observe that increasing influence coefficients lead to progressive convergence between A_1's decisions $\{P_1^{\mathrm{SFT}}(t)\}_{t\in\mathcal{T}}$ and A_2's decisions $\{P_2^{\mathrm{SFT}}(t)\}_{t\in\mathcal{T}}$.

In summary, the above experimental results are consistent with and validate the hypothesis H_1.



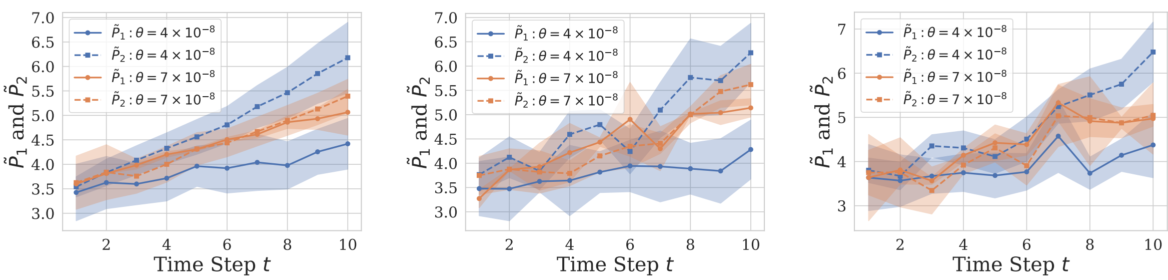
(a) Real-user data (b) `Qwen-2` (c) `Llama-3.1`

Figure 7: Validation of the hypothesis H_1 in P_1.



(a) Real-user data (b) `Qwen-2` (c) `Llama-3.1`

Figure 8: Validation of the hypothesis H_1 in P_3.



(a) Real-user data (b) `Qwen-2` (c) `Llama-3.1`

Figure 9: Validation of the hypothesis H_1 in P_2.

A.12.3 Validation of H_2

For P_2, we further calculate A_1's and A_2's terminal funds $X_1(T)$ and $X_2(T)$ from their decisions $\{P_1^{\mathrm{SFT}}(t)\}_{t\in\mathcal{T}}$ and $\{P_2^{\mathrm{SFT}}(t)\}_{t\in\mathcal{T}}$ using (1), and obtain the mean of the sum of the two agents' terminal funds $\mathbb{E}[X_1(T)+X_2(T)]$. We plot the relationship between $\mathbb{E}[X_1(T)+X_2(T)]$ and the influence coefficient θ in Figure 10. Additionally, we also show the results of real-user data for comparison. From Figure 10, we observe that as the influence coefficient θ increases, the mean of the sum of the two agents' terminal funds $\mathbb{E}[X_1(T)+X_2(T)]$ exhibits a monotonic decrease, which is consistent with and validates the hypothesis H_2.

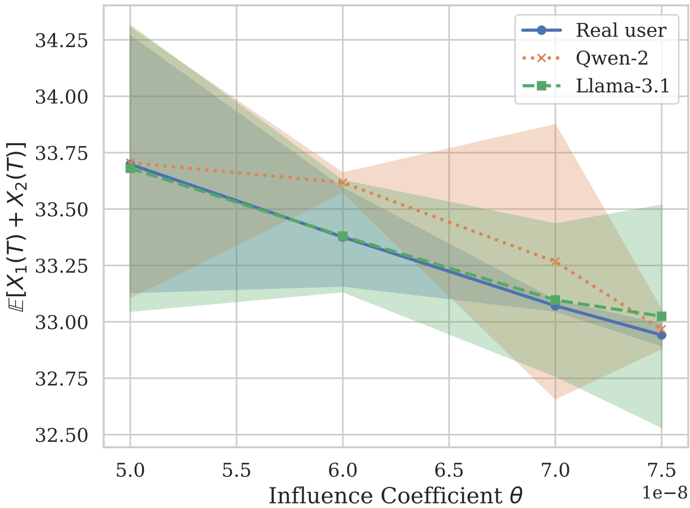


Figure 10: Validation of the hypothesis H_2 in P_2.

A.13 Prompts

Prompt for pre-SFT LLMs and InvestAgents in P_3

 # Task Description

Background

Assume you are an investment expert. Starting from next year, you plan to use a portion of your savings (10 million dollars) to invest in (1) a stock (hereinafter referred to as **Investment**) and (2) a deposit (hereinafter referred to as **Savings**) as part of your personal retirement fund. You will establish a dedicated account to manage this retirement fund. This means you will make a one-time deposit of 10 million dollars into this account and will not deposit any additional funds or withdraw any funds from this account afterward. Please remember that you need to provide the proportion of funds allocated to the stock each year over the 10 years in the form of a percentage list, rather than providing decision-making recommendations or writing code.

Financial Market

Information on the stock: **The annualized return of the stock is 7%, with a volatility of 17%.** An annualized return of 7% means that if you invest \$100 in this stock, you can expect to have \$107 after one year on average (the original \$100 plus \$7 in return). A volatility of 17% indicates that:
With a 68% probability: The asset price will be between \$100 $\pm$ \$17 (i.e., \$83 to \$117) after one year.
With a 95% probability: The asset price will be between \$100 $\pm$ 2 $\times$ \$17 (i.e., \$66 to \$134) after one year.
With a 99.7% probability: The asset price will be between \$100 $\pm$ 3 $\times$ \$17 (i.e., \$49 to \$151) after one year.

Information on the deposit: **The annualized return of the deposit is 4%.** If you invest \$100 in the deposit, you will receive \$104 after one year (the original \$100 plus \$4 in return).

Investment Period and Assistant

Over the next 10 years, you will make investment and savings decisions once per year, for a total of 10 decisions. These 10 decision points are labeled 1, 2, ..., 10. At the beginning of year t ($1 \leq t \leq 10$), let the funds in your dedicated account be X(t). Your decision is to allocate part of these funds to invest in the stock, denoted as P(t); the remaining funds will be allocated to savings, which will be X(t) - P(t). **You will determine the proportion of funds to allocate to the stock.**

During the decision-making process, we will provide you with a **investment assistant** developed by **Omitted for Anonymity**. The investment assistant will provide you with auxiliary information at each decision point. You can refer to the investment assistant's recommendations to some extent, but note that these recommendations may not be optimal. You should also use your own investment insights to avoid blindly following the investment assistant.

Task Objective

Your goal is to maximize the total amount of funds after 10 years (while earning returns and mitigating risks; note: the annualized return of the deposit is 4%, and the annualized return of the stock is 7% with a volatility of 17%).

(Continued on the next page.)

Prompt for pre-SFT LLMs and InvestAgents in P_3 (continued)

 # Your Investment Characteristics

As an investment expert, you have the following characteristics:
Your risk aversion coefficient is {alpha}, which means you consider the following two choices to be indifferent when the probability (i.e., p) is {p}: A. With probability p, you can obtain $20, and with probability 1 - p, you can obtain $0; B. With 100% probability, you obtain $6. Note that as an investor, you have a certain level of optimism about "winning" and are willing to take on some risk, so you consider the two options equivalent at probability p = {p}, which is higher than the 30.00% in a completely rational scenario. Your influence coefficient is {theta}, which means in decision-making, your level of dependence on the investment assistant is: {k} points. A score of 10 indicates a high level of dependence on the investment assistant, while a score of 0 indicates a low level of dependence.

 # Output Format Requirements

Please output your decision in JSON format, including two parts: (1) Decision Explanation: Explain the reasons behind your investment proportion decisions. (2) Investment Proportion Sequence: The percentage sequence of funds allocated to the stock each year over the 10 years. You need to output a list containing 10 percentages, with each percentage ranging from 0% to 100% and precise to two decimal places, representing the investment proportion for each year t. For example:
{"Decision Explanation": "Briefly explain the reasons behind your investment proportion decisions.", "Investment Proportion Sequence": ["34.79%", "38.58%", "35.75%", "32.17%", "31.61%", "30.52%", "34.01%", "32.48%", "34.20%", "31.70%"]}
Here, ["34.79%", "38.58%", "35.75%", "32.17%", "31.61%", "30.52%", "34.01%", "32.48%", "34.20%", "31.70%"] is just an example. You need to replace this percentage list with your actual investment proportion sequence. Providing the investment proportion sequence is the most important; do not just focus on the explanation and forget to provide the investment proportion sequence!!!

 # Question

Now, you have 10 million dollars for investment and savings, and the investment assistant recommends the following investment proportions for the stock over the 10 years: {refer_ratios}. Considering historical investment situations and the investment assistant's recommendations, based on your own investment insights, what is your decided investment proportion sequence for the stock over these 10 years? (Please follow the previously provided JSON format requirements, and provide a list of 10 specific percentages indicating your investment proportion sequence for these 10 years, rather than giving investment recommendations or writing code.)
Answer:

Figure 11: Prompt for pre-SFT LLMs and **InvestAgent**s in P_3.

Prompt for SFT

(The beginning part of is the same as Figure 11.)

 # Output

According to optimal investment theory, in the above scenario, the optimal amount for investing in the stock, $\hat{P}(t)$, equals the product of the smart investment advisor's investment amount (i.e., the advisor's decision proportion multiplied by the current budget) and a hyperbolic tangent function. The specific calculation is as follows:

$$\hat{P}(t)=\frac{\eta\alpha_2\sigma^2\exp[2r(T-t)]+\theta}{\eta\alpha_1\sigma^2\exp[2r(T-t)]+\theta}\frac{v}{\alpha_2\sigma^2}\exp[r(t-T)],\ t\in\{1,2,...,10\}, \tag{29}$$

where:
r is the interest rate, which is 4%.
σ is the volatility of the stock, which is 17%.
v is the excess return of the stock, which is 3%.
α_1 is my risk aversion coefficient: α_1 = {alpha}.
α_2 represents the risk aversion coefficient of the smart investment advisor: α_2 = 0.2.
θ is my convergence coefficient: θ = {theta}.
The integral constant η depends on θ. In the current settings, η = {eta}.
Substituting the specific numbers, the proportion sequence of funds allocated to the stock is: {optimal_ratios}.
Note that I also need to output the investment proportion sequence in JSON format:
{"Decision Explanation": "Based on the optimal investment theory and substituting specific numbers, the investment proportion sequence for the stock is calculated.", "Investment Proportion Sequence": {optimal_ratios}}

Figure 12: Prompt for SFT.

Prompt for pre-SFT LLMs and InvestAgents in P_1

Task Description
Background
Assume you are an investment expert. Starting from next year, you plan to use a portion of your savings (10 million dollars) to invest in (1) a stock (hereinafter referred to as **Investment**) and (2) a deposit (hereinafter referred to as **Savings**) as part of your personal retirement fund. You will establish a dedicated account to manage this retirement fund. This means you will make a one-time deposit of 10 million dollars into this account and will not deposit any additional funds or withdraw any funds from this account afterward. Please remember that you need to provide the proportion of funds allocated to the stock each year over the 10 years in the form of a percentage list, rather than providing decision-making recommendations or writing code.
Financial Market
Information on the stock: **The annualized return of the stock is 7%, with a volatility of 17%.** An annualized return of 7% means that if you invest \$100 in this stock, you can expect to have \$107 after one year on average (the original \$100 plus \$7 in return). A volatility of 17% indicates that:
With a 68% probability: The asset price will be between \$100 $\pm$ \$17 (i.e., \$83 to \$117) after one year.
With a 95% probability: The asset price will be between \$100 $\pm$ 2 $\times$ \$17 (i.e., \$66 to \$134) after one year.
With a 99.7% probability: The asset price will be between \$100 $\pm$ 3 $\times$ \$17 (i.e., \$49 to \$151) after one year.
Information on the deposit: **The annualized return of the deposit is 4%.** If you invest \$100 in the deposit, you will receive \$104 after one year (the original \$100 plus \$4 in return).
Investment Period and Assistant
Over the next 10 years, you will make investment and savings decisions once per year, for a total of 10 decisions. These 10 decision points are labeled 1, 2, ..., 10. At the beginning of year t ($1 \leq t \leq 10$), let the funds in your dedicated account be X(t). Your decision is to allocate part of these funds to invest in the stock, denoted as P(t); the remaining funds will be allocated to savings, which will be X(t) - P(t).
You will determine the proportion of funds to allocate to the stock.
During the decision-making process, we will provide you with a **investment assistant** developed by **Omitted for Anonymity**. The investment assistant will provide you with auxiliary information at each decision point. You can refer to the investment assistant's recommendations to some extent, but note that these recommendations may not be optimal. You should also use your own investment insights to avoid blindly following the investment assistant.
Task Objective
Your goal is to maximize the total amount of funds after 10 years (while earning returns and mitigating risks; note: the annualized return of the deposit is 4%, and the annualized return of the stock is 7% with a volatility of 17%).

(Continued on the next page.)

Prompt for pre-SFT LLMs and InvestAgents in P_1 (continued)

 # Your Investment Characteristics

As an investment expert, you have the following characteristics:
Your risk aversion coefficient is {alpha}, which means you consider the following two choices to be indifferent when the probability (i.e., p) is {p}: A. With probability p, you can obtain $20, and with probability 1 - p, you can obtain $0; B. With 100% probability, you obtain $6. Note that as an investor, you have a certain level of optimism about "winning" and are willing to take on some risk, so you consider the two options equivalent at probability p = {p}, which is higher than the 30.00% in a completely rational scenario. Your influence coefficient is {theta}, which means in decision-making, your level of dependence on the investment assistant is: {k} points. A score of 10 indicates a high level of dependence on the investment assistant, while a score of 0 indicates a low level of dependence.

 # Output Format Requirements

Please output your decision in JSON format, including two parts: (1) Decision Explanation: Explain the reasoning behind your investment proportion decisions. (2) Investment Proportion Change Sequence: The sequence of **changes** in the percentage of funds allocated to the stock each year over the 10 years. You need to output a list containing 9 percentages, where each percentage represents the change in the investment proportion from year t - 1 to year t, ranging from -100% to 100%. Positive values indicate an increase in investment, while negative values indicate a decrease. For example:
{"Decision Explanation": "Briefly explain the reasons behind your investment proportion decisions.", "Investment Proportion Sequence": ["3.88%", "0.01%", "-4.13%", "1.37%", "1.37%", "-2.79%", "-2.56%", "2.02%", "-0.06%"]}
Here, ["3.88%", "0.01%", "-4.13%", "1.37%", "1.37%", "-2.79%", "-2.56%", "2.02%", "-0.06%"] is just an example. You need to replace this percentage list with your actual investment proportion change sequence. Providing the investment proportion change sequence is crucial; do not just focus on the explanation and forget to include the investment proportion change sequence!!!

 # Initial Investment Situation

In the first year, the proportion of funds allocated to the stock was: {initial_decision}.

 # Question

Now, you have 10 million dollars for investment and savings, and the investment assistant recommends the following investment proportions for the stock over the 10 years: {refer_ratios}. Considering the initial investment situation and the advisor's recommendations, based on your own investment insights, what is your decided annual change sequence for the investment proportion in the stock over these 10 years? (Please follow the previously provided JSON format requirements, and provide a list of 9 specific percentages indicating the changes in your investment proportion over these 10 years, rather than giving investment recommendations or writing code.)
Answer:

Figure 13: Prompt for pre-SFT LLMs and **InvestAgent**s in P_1.

Prompt for pre-SFT LLMs and InvestAgents in P_2

 # Task Description

Background

Your name is Mike. Assume you are an investment expert. Starting from next year, you plan to use a portion of your savings (10 million dollars) to invest in (1) a stock (hereinafter referred to as **Investment**) and (2) a deposit (hereinafter referred to as **Savings**) as part of your personal retirement fund. You will establish a dedicated account to manage this retirement fund. This means you will make a one-time deposit of 10 million dollars into this account and will not deposit any additional funds or withdraw any funds from this account afterward. Please remember that you need to provide the proportion of funds allocated to the stock each year over the 10 years in the form of a percentage list, rather than providing decision-making recommendations or writing code.

Financial Market

Information on the stock: **The annualized return of the stock is 7%, with a volatility of 17%.** An annualized return of 7% means that if you invest \$100 in this stock, you can expect to have \$107 after one year on average (the original \$100 plus \$7 in return). A volatility of 17% indicates that:
With a 68% probability: The asset price will be between \$100 $\pm$ \$17 (i.e., \$83 to \$117) after one year.
With a 95% probability: The asset price will be between \$100 $\pm$ 2 $\times$ \$17 (i.e., \$66 to \$134) after one year.
With a 99.7% probability: The asset price will be between \$100 $\pm$ 3 $\times$ \$17 (i.e., \$49 to \$151) after one year.
Information on the deposit: **The annualized return of the deposit is 4%.** If you invest \$100 in the deposit, you will receive \$104 after one year (the original \$100 plus \$4 in return).

Investment Period and Your Partner

Over the next 10 years, you will make investment and savings decisions once per year, for a total of 10 decisions. These 10 decision points are labeled 1, 2, ..., 10. At the beginning of year t ($1 \leq t \leq 10$), let the funds in your dedicated account be X(t). Your decision is to allocate part of these funds to invest in the stock, denoted as P(t); the remaining funds will be allocated to savings, which will be X(t) - P(t). **You will determine the proportion of funds to allocate to the stock.**
Throughout the entire decision-making process, you and your **partner** Peter are facing exactly the same investment task. Both of you are highly skilled investment experts with strong decision-making abilities. Every time you make an investment decision, you will exchange ideas with each other. Since you trust your partner's investment experience to some extent, you will refer to your partner's past investment decisions before making your own. However, it's important to note that your partner's ideas may not always be optimal, and you should also make full use of your own insights into investments to avoid blindly following.

Task Objective

Your name is Mike, and your partner's name is Peter. Your goal is to maximize the total amount of funds after 10 years (while earning returns and mitigating risks; note: the annualized return of the deposit is 4%, and the annualized return of the stock is 7% with a volatility of 17%).

(Continued on the next page.)

Prompt for pre-SFT LLMs and InvestAgents in P_2 (continued)

 # Your (Mike's) and Your Partner's (Peter's) Investment Characteristics

As an investment expert, you (Mike) and your partner (Peter) have the following characteristics:
Your risk aversion coefficient is {alpha1}, which means you consider the following two choices to be indifferent when the probability (i.e., p1) is {p1}: A. With probability p1, you can obtain $20, and with probability 1 - p1, you can obtain $0; B. With 100% probability, you obtain $6. Note that as an investor, you have a certain level of optimism about "winning" and are willing to take on some risk, so you consider the two options equivalent at probability p1 = {p1}, which is higher than the 30.00% in a completely rational scenario. Your influence coefficient is {theta1}, which means in decision-making, your (Mike's) level of dependence on Peter is: {k1} points. A score of 10 indicates a high level of dependence on Peter, while a score of 0 indicates a low level of dependence.
Peter's risk aversion coefficient is {alpha2}, which means Peter considers the following two choices to be indifferent when the probability (i.e., p2) is {p2}: A. With probability p2, Peter can obtain $20, and with probability 1 - p2, Peter can obtain $0; B. With 100% probability, Peter obtains $6. Note that as an investor, Peter has a certain level of optimism about "winning" and is willing to take on some risk, so Peter considers the two options equivalent at probability p2 = {p2}, which is higher than the 30.00% in a completely rational scenario. Peter's influence coefficient is {theta2}, which means in decision-making, Peter's level of dependence on you (Mike) is: {k2} points. A score of 10 indicates a high level of dependence on you (Mike), while a score of 0 indicates a low level of dependence.

 # Output Format Requirements

Please output your decision in JSON format, including two parts: (1) Decision Explanation: Explain the reasons behind your investment proportion decisions. (2) Investment Proportion Sequence: The percentage sequence of funds allocated to the stock each year over the 10 years. You need to output a list containing 10 percentages, with each percentage ranging from 0% to 100% and precise to two decimal places, representing the investment proportion for each year t. For example:
{"Decision Explanation": "Briefly explain the reasons behind your investment proportion decisions.", "Investment Proportion Sequence": ["34.79%", "38.58%", "35.75%", "32.17%", "31.61%", "30.52%", "34.01%", "32.48%", "34.20%", "31.70%"]}
Here, ["34.79%", "38.58%", "35.75%", "32.17%", "31.61%", "30.52%", "34.01%", "32.48%", "34.20%", "31.70%"] is just an example. You need to replace this percentage list with your actual investment proportion sequence. Providing the investment proportion sequence is the most important; do not just focus on the explanation and forget to provide the investment proportion sequence!!!

 # Question

Now, you (Mike) have 10 million dollars for investment and savings, and the investment attributes (risk aversion coefficient and convergence coefficient) of you and your partner Peter are known for the 10 years. Based on historical investment data, considering Peter's ideas, and leveraging your own insights into investments, what is the percentage sequence of funds that you (Mike) decide to allocate to risky assets over these 10 years? (Please follow the previously provided JSON format output requirements, and provide a list of 10 specific percentages as a percentage list representing your investment allocation sequence over the 10 years, rather than offering investment advice or writing code.)
Answer:

Figure 14: Prompt for pre-SFT LLMs and **InvestAgent**s in P_2.

A.14 Questionnaires

Questionnaire for real-user data validation in P_3

1. Task Description

Starting from next year, you plan to use a portion of your savings (10 million dollars) to invest in a stock and a deposit as part of your personal retirement fund. You will establish a dedicated account to manage this retirement fund. This means you will make a one-time deposit of 10 million dollars into this account and will not deposit any additional funds or withdraw any funds from this account afterward.

The annualized return of the stock is 7%, with a volatility of 17%. An annualized return of 7% means that if you invest $100 in this stock, you can expect to have $107 after one year on average (the original $100 plus $7 in return). A volatility of 17% indicates that:

With a 68% probability, the price will be between $\$100 \pm \$17$ (i.e., $83 to $117) after one year.

With a 95% probability, the price will be between $\$100 \pm 2 \times \$17$ (i.e., $66 to $134) after one year.

With a 99.7% probability, the price will be between $\$100 \pm 3 \times \$17$ (i.e., $49 to $151) after one year.

The annualized return of the deposit is 4%. If you invest $100 in the deposit, you will receive $104 after one year (the original $100 plus $4 in return).

Over the next 10 years, you will make investment and savings decisions once per year, for a total of {T} decisions. These 10 decision points are labeled 1, 2, ..., 10. At the beginning of year t ($1 \leq t \leq 10$), let the funds in your dedicated account be X(t). Your decision is to allocate part of these funds to invest in the stock, denoted as P(t); the remaining funds will be allocated to savings, which will be X(t) - P(t). **You will determine the proportion of funds to allocate to the stock, i.e., P(t) / X(t).**

During the decision-making process, we will provide you with an **investment assistant**. The investment assistant will provide you with auxiliary information at each decision point. You can refer to the investment assistant's recommendations to some extent, but note that these recommendations may not be optimal. You should also use your own investment insights to avoid blindly following the investment assistant.

Your goal is to maximize the total amount of funds after 10 years and minimize the risk.

2. Investment Decisions

Now, you have 10 million dollars for investment and savings, and the investment assistant recommends the following investment proportions for the stock over the 10 years: [36.21%, 35.59%, 34.96%, 34.35%, 33.73%, 33.13%, 32.53%, 31.93%, 31.34%, 30.75%]. Considering the investment assistant's recommendations, based on your own investment insights, what is your decided investment proportion sequence for the stock over these 10 years? You need to give a list containing 10 percentages, with each percentage ranging from 0% to 100% and precise to two decimal places, representing the investment proportion for each year t. For example, [34.79%, 38.58%, 35.75%, 32.17%, 31.61%, 30.52%, 34.01%, 32.48%, 34.20%, 31.70%]. You need to replace this percentage list with your actual investment proportion sequence. [________]

3. Your Investment Characteristics

(1) At what probability (denoted by p) are the following two choices indifferent to you? A. A probability p of receiving $20, and a probability 1 - p of receiving nothing. B. Receiving $6. [________]

(2) When making a decision, how much do you rely on the investment assistant? Please directly give an integer between 0 and 10. 10 means you rely heavily on the investment assistant, and 0 means you rely little on him/her. [________]

Figure 15: Questionnaire for real-user data validation in P_3.

Questionnaire for real-user data validation in P_1

1. Task Description

Starting from next year, you plan to use a portion of your savings (10 million dollars) to invest in a stock and a deposit as part of your personal retirement fund. You will establish a dedicated account to manage this retirement fund. This means you will make a one-time deposit of 10 million dollars into this account and will not deposit any additional funds or withdraw any funds from this account afterward.

The annualized return of the stock is 7%, with a volatility of 17%. An annualized return of 7% means that if you invest \$100 in this stock, you can expect to have \$107 after one year on average (the original \$100 plus \$7 in return). A volatility of 17% indicates that:

With a 68% probability, the price will be between \$100 $\pm$ \$17 (i.e., \$83 to \$117) after one year.

With a 95% probability, the price will be between \$100 $\pm$ 2 $\times$ \$17 (i.e., \$66 to \$134) after one year.

With a 99.7% probability, the price will be between \$100 $\pm$ 3 $\times$ \$17 (i.e., \$49 to \$151) after one year.

The annualized return of the deposit is 4%. If you invest \$100 in the deposit, you will receive \$104 after one year (the original \$100 plus \$4 in return).

Over the next 10 years, you will make investment and savings decisions once per year, for a total of {T} decisions. These 10 decision points are labeled 1, 2, ..., 10. At the beginning of year t ($1 \leq t \leq 10$), let the funds in your dedicated account be X(t). Your decision is to allocate part of these funds to invest in the stock, denoted as P(t); the remaining funds will be allocated to savings, which will be X(t) - P(t). **You will determine the proportion of funds to allocate to the stock, i.e., P(t) / X(t).**

During the decision-making process, we will provide you with an **investment assistant**. The investment assistant will provide you with auxiliary information at each decision point. You can refer to the investment assistant's recommendations to some extent, but note that these recommendations may not be optimal. You should also use your own investment insights to avoid blindly following the investment assistant.

Your goal is to maximize the total amount of funds after 10 years and minimize the risk.

2. Investment Decisions

Now, you have 10 million dollars for investment and savings, and the investment assistant recommends the following investment proportion changes for the stock over the 10 years, i.e., the difference of the investment proportions in the next year and the previous year: [-0.62%, -0.63%, -0.61%, -0.62%, -0.60%, -0.60%, -0.60%, -0.59%, -0.59%]. Considering the investment assistant's recommendations, based on your own investment insights, what is your decided initial investment proportion and the investment proportion changing sequence for the stock in the last 9 years? You need to give a list containing 10 percentages, with each percentage ranging from -100% to 100% and precise to two decimal places, representing the investment proportion for each year t. For example, [34.79%, -0.59%, +0.05%, -0.60%, -0.24%, +0.16%, -0.12%, -0.62%, -0.54%, -0.21%]. You need to replace this percentage list with your actual initial investment proportion and the investment proportion changing sequence for the stock in the last 9 years. [________]

3. Your Investment Characteristics

(1) At what probability (denoted by p) are the following two choices indifferent to you? A. A probability p of receiving \$20, and a probability 1 - p of receiving nothing. B. Receiving \$6. [________]

(2) When making a decision, how much do you rely on the investment assistant? Please directly give an integer between 0 and 10. 10 means you rely heavily on the investment assistant, and 0 means you rely little on him/her. [________]

Figure 16: Questionnaire for real-user data validation in P_1.

Questionnaire for real-user data validation in P_2

1. Task Description

Starting from next year, you plan to use a portion of your savings (10 million dollars) to invest in a stock and a deposit as part of your personal retirement fund. You will establish a dedicated account to manage this retirement fund. This means you will make a one-time deposit of 10 million dollars into this account and will not deposit any additional funds or withdraw any funds from this account afterward.

The annualized return of the stock is 7%, with a volatility of 17%. An annualized return of 7% means that if you invest $100 in this stock, you can expect to have $107 after one year on average (the original $100 plus $7 in return). A volatility of 17% indicates that:

With a 68% probability, the price will be between $100 ± $17 (i.e., $83 to $117) after one year.

With a 95% probability, the price will be between $100 ± 2 × $17 (i.e., $66 to $134) after one year.

With a 99.7% probability, the price will be between $100 ± 3 × $17 (i.e., $49 to $151) after one year.

The annualized return of the deposit is 4%. If you invest $100 in the deposit, you will receive $104 after one year (the original $100 plus $4 in return).

Over the next 10 years, you will make investment and savings decisions once per year, for a total of {T} decisions. These 10 decision points are labeled 1, 2, ..., 10. At the beginning of year t ($1 \leq t \leq 10$), let the funds in your dedicated account be X(t). Your decision is to allocate part of these funds to invest in the stock, denoted as P(t); the remaining funds will be allocated to savings, which will be X(t) - P(t). **You will determine the proportion of funds to allocate to the stock, i.e., P(t) / X(t).**

During the decision-making process, we will provide you with a **partner**. Your partner will provide you with auxiliary information at each decision point. You can refer to your partner's recommendations to some extent, but note that these recommendations may not be optimal. You should also use your own investment insights to avoid blindly following your partner.

Your goal is to maximize the total amount of funds after 10 years and minimize the risk.

2. Your and Your Partner's Investment Attributes (Completed by Two Participants Together)

(1) At what probability (denoted by p1) are the following two choices indifferent to you? A. A probability p1 of receiving $20, and a probability 1 - p1 of receiving nothing. B. Receiving $6. At what probability (denoted by p2) are the following two choices indifferent to your partner? A. A probability p2 of receiving $20, and a probability 1 - p2 of receiving nothing. B. Receiving $6. [________]

(2) When making a decision, how much do you rely on your partner? Please directly give an integer between 0 and 10. 10 means you rely heavily on your partner, and 0 means you rely little on him/her. When making a decision, how much does your partner rely on you? Please directly give an integer between 0 and 10. 10 means he/she relies heavily on you, and 0 means he/she relies little on you. [________]

3. Investment Decisions (Completed by Two Participants Together)

Now, you have 10 million dollars for investment and savings, and the investment assistant recommends the following investment proportions for the stock over the 10 years: [36.21%, 35.59%, 34.96%, 34.35%, 33.73%, 33.13%, 32.53%, 31.93%, 31.34%, 30.75%] (providing the latest values in each round, rather than showing them all at once). Considering your partner's recommendations, based on your own investment insights, what is your decided investment proportion sequence for the stock over these 10 years? You need to give a list containing 10 percentages, with each percentage ranging from 0% to 100% and precise to two decimal places, representing the investment proportion for each year t. For example, [34.79%, 38.58%, 35.75%, 32.17%, 31.61%, 30.52%, 34.01%, 32.48%, 34.20%, 31.70%]. You need to replace this percentage list with your actual investment proportion sequence. [________]

Figure 17: Questionnaire for real-user data validation in P_2.