Research proposal

Applicant's name:

Peng Yifeng

Proposed topic:

New behavioral biases with the application of large language model in financial decision-making.

Abstract:

The rapid development of fintech has brought about the widespread application of large language models (LLMS) in financial decision making. However, effectively integrating the interactions between LLMS and human decision makers, as well as managing possible cognitive biases, is one of the key challenges to achieving optimal decision making. This research program aims to investigate the behavioral bias of LLM in financial decision making, with a particular focus on human-machine alignment (HMA), and seeks to learn and apply strategies for the detection and correction of cognitive bias from a behavioral finance perspective. This study intends to use a hybrid approach research design, combining human-computer interaction experiments, surveys, and data analysis, to evaluate the biases that occur in the LLM

simulation of financial decision making and propose strategies for improvement. The research is expected to provide strategic guidance to the financial services industry, promote the effectiveness and accountability of LLM in financial decision support, and contribute new theoretical and practical insights to the intersection of artificial intelligence and behavioral finance.

Research Background:

Generative artificial intelligence has been heralded as a technological revolution within the information era, with large language models (LLMs) suffusing multiple industries to augment human efficiency. Within the rapidly evolving domain of financial technology, generative artificial intelligence, coupled with LLM-based financial advisory systems, is assuming a pivotal role in the realm of financial decision-making support[1][2]. These models are instrumental in providing nuanced insights into the intricate dynamics of financial markets by leveraging their capabilities in the analysis and processing of extensive datasets. An increasing number of financial institutions, including governmental entities, are adopting LLMs to enhance the efficiency and precision of trading strategies, risk assessments, and customer service[3][4].

Currently, the parameters and training data volumes of commercially available large language models (LLMs) are already vast, and their scale is expected to increase further in the foreseeable future. While the code that structures these complex models and the training methods used can be straightforward, the models generated posttraining are highly abstruse and intricate. Users are often left to regard the models as a "black box," accepting certain inputs and then outputting relevant content to meet the needs of the users. Despite considerable research efforts in the field of model interpretability, the process by which these models generate outputs with each iteration remains largely inexplicable even to the scientists who train them. The "black box" nature of artificial intelligence generative content (AIGC) raises concerns about transparency, accountability, and ethical implications[5][6]. Relevant studies to discuss and analyze such alignment problems show that there are extensive and difficult to detect biases in the understanding of human instructions by large language models[7][8].

A dialectical perspective towards AIGC would become increasingly necessary with the ensuing blossom of LLMs. Recent studies have shown that LLM may inadvertently amplify human cognitive biases in financial decision-making[16] or create new problems due to algorithms' own biases[17][18]. E.g., LLMS may over-rely on historical data patterns when forecasting financial markets[19], while ignoring real-time fluctuations in market conditions. In addition, human users may over-reliance on the recommendations generated by LLM, leading to the weakening of critical thinking in the decision-making process[20]. Besides, when the recommendations generated by LLM are deviated from the needs or even contrary to common sense, the ethical and moral issues that may lead to disastrous consequences are difficult to be summarized[18][20].

With the rapid advancements in AI technology, human contemplation of behavior has expanded from natural biological entities to artificial agents. As noted in the research by Iyad et al.[21], AI is becoming an integral component of human society, where the behavior of intelligent agents extends beyond the fixed frameworks set by their creators. The study of these behavior patterns is instrumental in helping humans understand and predict the "deviant" actions of machines. Drawing from four-dimensional view of ethology[22], machine behavior fundamentally differs from that of animals and humans. It is essential that the study of machine behavior avoids excessive anthropomorphism nor zoomorphism. The ultimate goal of researching machine behavior should be to enhance social welfare by aiding and augmenting human decision-making capabilities.

As AI systems, including LLMs, are increasingly integrated into financial advisory and decision-support roles, their interaction with human cognitive biases becomes a pivotal area of study. The crux of the human-machine alignment problem across various domains stems from the discordance between the intentions of human users and the behaviors exhibited by the agents. The core findings of behavioral finance reveal systematic biases in the way human decision-makers face financial decisions. For example, the forward-looking theory of Tversky and Kahneman[9][25] explains how people systematically deviate from the predictions of expected utility theory in the face of uncertainty. These biases, such as overconfidence, emotional influence, and anchoring effects, not only affect individual investors, but are also reflected in institutional investment decisions[10][11]. In order to tame the innate perceptual cognitive emotions in the human mind[12], a school of research exploring how to integrate the understanding of these behavioral biases into the LLM to optimize their performance in real-world decision-making is gaining attention[13][14][15].

To address these issues, this research study presents an innovative research topic: how to guide the design and application of LLMS with professional finance backgrounds through a deep understanding of the mechanisms of cognitive bias in behavioral finance, achieve effective human-machine alignment (HMA), and optimize the financial decision-making process. Human-machine alignment refers to the degree to which the behavior of an AI system aligns with the expectations and goals of a human user[20][23][24]. In the financial field, this means that LLM can provide customized decision support based on considering users' preferences, risk tolerance and behavioral characteristics[20].

In summary, this study aims to fill in the gaps in the current research on the cognitive bias treatment and human-machine alignment optimization of LLM in financial decision making in behavioral finance and human-machine alignment. Through a combination of empirical research and interdisciplinary theoretical analysis, we are committed to providing more accurate, transparent and ethical AI decision-making tools for financial services.

Research Aim:

The primary objective of this study is to examine the manifestation of cognitive biases within LLMs as they pertain to financial decision-making processes. The investigation is further aimed at devising and implementing strategies to rectify these biases, thereby enhancing the utility and accuracy of LLMs in providing decision support within the financial sector. The research aims to make a significant theoretical contribution to the understanding of cognitive biases in artificial intelligence, while also offering practical insights for the refinement of LLMs applications in finance. Specifically, the objectives of this study can be divided into the following aspects:

- **Understanding Behavioral Bias Mechanisms:**

  This research will delve into the mechanisms of behavioral biases in financial decisions, drawing on established theories and empirical studies to identify prevalent cognitive biases such as overconfidence, anchoring effects, and loss aversion.

- **Evaluating LLMs in Financial Decision-Making:**

  The study will assess the efficacy of current LLMs in interpreting financial decisions, with a focus on their ability to process and mitigate cognitive biases in financial information.

- **Enhancing Human-Machine Alignment:**

  By leveraging recent advancements in reinforcement learning and AI behavior, this research will explore strategies to enhance the alignment between human users and LLMs, aiming to refine model training and integrate bias detection and correction mechanisms.

- **Advancing Rational Financial Decision-Making:**

  The ultimate goal is to foster a more rational approach to financial decision-making by optimizing the management of behavioral biases within LLMs, thereby enhancing the efficiency and equity of financial markets.

Literature Review:

Behavioral finance and cognitive bias

Behavioral finance is a discipline which studies the influence of investors' psychological behavior on financial markets and decision making. It stems from the limitations of traditional finance, particularly in understanding market anomalies and irrational investor behavior. Different from the traditional finance assumption that market participants are always rational[35], behavioral finance believes that people are often affected by various cognitive biases when making financial decisions. The core of financial decision-making is the balance between risk assessment and expected return. However, traditional financial theories, such as the efficient market hypothesis [35], believe that investors are rational, but in reality, investors' behavior often deviates from rationality[36]. Identifying and understanding the causes of these deviations, especially the role of cognitive bias, is important for financial decision-making.

Kahneman and Tversky's prospect theory [9], one of the most influential theories in behavioral finance, challenges expected utility theory and proposes the nonlinear preferences that people exhibit when faced with gains and losses. They found that people are more sensitive to potential losses relative to expected absolute gains, which is known as "loss aversion" [39]. In addition, other biases like the representativeness bias[26], which explains how people evaluate probabilities based on preconceived judgments rather than statistical evidence; anchoring effect [9]: People tend to rely on first impressions or initial information in the face of uncertainty; And the availability heuristic[27], which describes investors' tendency to over-rely on information that comes to mind easily.

Subsequent research has continued to expand the range of cognitive biases to include overconfidence, in which individuals are overly confident in the correctness of their own judgments [37]; And confirmation bias, the tendency of individuals to seek, interpret, and remember information to confirm their preconceptions. [38]

These biases are particularly pronounced in financial markets. For example, overconfidence leads individual investors to overtrade, which often reduces their investment returns. [37] However, the study of Shefrin and Statman shows that investors are often unwilling to admit investment mistakes and cut losses in time, which is the behavioral embodiment of loss aversion[39].

Individual investors are often affected by a variety of cognitive biases when making financial decisions. For example, confirmation bias causes investors to tend to seek out and remember information that supports their preconceived notions. [40] This can lead retail investors to ignore important market information and make sub-optimal investment decisions. Mood swings are also a key factor in individual financial decisions. The research of Lerner, Li, Valdesolo and Kassam shows that different emotional states can significantly affect people's risk preference, such as anger tends to increase risk taking, while fear reduces risk taking [41].

In the company's financial decision-making, the cognitive bias of the management will also have an impact on the company's investment, financing and dividend policies. For example, overconfidence bias may lead managers to overestimate their project management capabilities or market prospects, leading to excessive investment [42]. In addition, agency problems are also associated with cognitive bias in corporate financial decision-making. The inconsistency of interests between managers and shareholders may lead managers to make decisions based on their own preferences rather than the best interests of shareholders, such as empire building and risk aversion [43].Group behaviors in the market, such as herding and bubble formation, are closely related to individual cognitive biases. Investors' behavior in groups may be driven by social pressures and the tendency to imitate others, rather than guided by independent information analysis. This behavior can cause market prices to deviate from their fundamental values, increasing the volatility of the market [44].

In financial markets, the game between different market participants is also affected by cognitive bias. Game theory studies have shown that even in the case of the rational player hypothesis, market participants may make sub-optimal decisions due to misunderstanding of the opponent's behavior or incorrect expectations [45]. This sub-optimal behavior may be more prevalent when participants are subject to cognitive bias. To reduce the impact of cognitive bias on financial decision-making, education and training can play a key role. Through financial education, investors can better understand market dynamics and learn to identify and avoid common psychological pitfalls[46][47]. In addition, professional investment training and simulation can help investors and financial professionals improve the quality of their decisions and learn how to manage their own biases and emotional responses through practice.

Application of large language model in financial decision making

With the development of artificial intelligence technology, especially the progress of natural language processing (NLP), large language models such as OpenAI's GPT series have begun to be applied in the financial field, including market analysis, risk assessment, investment strategy formulation, etc. [1][2][3][4][19]. By understanding and generating natural language, these models can extract valuable information from large amounts of textual data to help financial professionals make more informed decisions.

One of the critical applications of LLMs in finance is market analysis. By processing news articles, financial reports, and social media posts, LLMs can perform sentiment analysis, gauging the market's emotional tone. This information is crucial for predicting market movements, as investor sentiment is often a leading indicator of market trends[2][19][48]. For instance, LLMs can detect shifts in sentiment that may precede changes in stock prices, providing investors with valuable lead time to adjust their strategies.

LLMs also play a significant role in risk management by identifying potential risks in financial documents and communications. Their ability to understand context and nuance allows them to detect subtle signals that could indicate fraudulent activity or emerging risks that are not evident through traditional analysis methods[2][19].

In customer service, LLMs can automate responses to client inquiries, provide personalized financial advice, and streamline the customer experience. This application not only increases efficiency but also allows for the handling of a larger volume of customer interactions, which can lead to improved customer satisfaction and loyalty[2][49].

Man-machine alignment and correction of cognitive bias

The human-machine alignment problem concerns how to align the goals of an AI system with human values and goals [错误!未找到引用源。]. In the context of financial decision-making, this means the need to develop large language models that can understand human cognitive biases and take steps to correct them. At present, there is no universally accepted standard for measuring alignment.

Researchers such as Wang et al summarize the alignment problem as RICE: Robustness meaning that the stability of artificial intelligence systems needs to be ensured in various environments; Interpretability means that the operation and decision-making process of AI systems should be clear and understandable. Controllability: An AI system should operate under human direction and control; Ethicality states that AI systems should adhere to social norms and universal values. These four principles guide the alignment of AI systems with human intentions and values. They are not the end goal in themselves, but serve the intermediate goal of alignment.[20]

One possible approach is to train models to recognize bias and error information in text and then adjust their output to reduce the impact of bias. For example, models can be trained to recognize and adjust predictions based on overconfidence or loss aversion. In addition, models can be trained with manually labeled datasets that are specifically designed to reflect and correct for specific cognitive biases[50].

However, there are challenges to achieving effective human-machine alignment. First, to accurately identify and correct cognitive biases requires a deep understanding of the psychological underpinnings of these biases, which may be beyond the capabilities of current large language models[20]. Second, even if the model is able to identify the bias, how to correct it without compromising the model's performance is an open question.

Research Method/methodology:

The research methodology will focus on assessing cognitive biases in how humans interact with high-level large language models, such as OpenAI's GPT-3.5 or later, and exploring how these biases might affect their financial decisions.

Experimental setting

The behavior of participants when interacting with a large language model is analyzed through controlled experiments in online platforms or laboratory Settings. The experiment will involve simulating financial decision situations in which participants need to process information provided by a large language model and make investment choices.

Controlled experiment:

Create an experimental environment where participants interact with a large language model on a simulated investment platform. In this environment, participants are provided with various financial information and asked to make investment decisions.

Situational simulation:

Design different financial scenarios, such as stock market investment, bond selection, or asset allocation, to assess different types of cognitive bias.

Human-machine conversation recording:

All conversations between participants and the large language model are recorded to facilitate subsequent analysis of cognitive biases and information processing patterns.

Types of data expected to be collected

Quantitative data:

investment selection, click-through rate, investment time and rate of return collected through the experimental platform.

Qualitative data:

Qualitative information extracted from human-machine conversations, such as participants' queries, reasons for decisions, and feedback.

Study participants

Individuals with some experience in financial markets are invited to participate in the experiment. The study will recruit participants through online financial communities and investment forums, and ensure that the sample is sufficiently diverse in age, gender, and occupational background.

Data analysis

Statistical testing:

Descriptive statistics are applied to analyze the distribution of investment decisions, and T-tests and ANOVA are used to compare decision differences across different scenarios.

Regression analysis:

Multiple regression analysis is used to assess the relationship between specific cognitive biases (such as overconfidence, anchoring effects, etc.) and investment choice.

Content analysis:

Conduct content analysis of human-computer conversations to identify patterns and themes associated with cognitive bias.

Research Limitations

Sample representation: Due to the use of convenience sampling, the sample may not be fully representative of all financial market participants.

Experimental situations: Situations simulated in a laboratory or online platform may not fully replicate the complexity and stress of the real world.

Ethical Considerations

Privacy protection: Ensure that participants' privacy is protected and no personally identifiable information is disclosed.

Informed consent: Ensure that all participants have provided informed consent and understand the purpose and process of the experiment.

This research approach provides a comprehensive framework for assessing cognitive bias in human interactions with large language models, as well as quantitative and qualitative data for studying its potential impact on financial decision-making processes. Through this approach, this study can begin to understand and quantify the influence of large language models in financial decision making.

Expected result:

Quantitative analysis of expected results

Correlation between cognitive biases and investment decisions: Expected regression analyses will show significant correlations between specific cognitive biases (e.g., overconfidence, anchoring effects, risk aversion, etc.) and investment choices. For example, studies may find that overconfident individuals tend to trade more frequently or choose riskier investment options.

Decision Differences in human-computer interaction:

It is expected that through T-test and ANOVA analysis, participants will exhibit different investment behaviors when interacting with and not interacting with the large language model. This may shed light on how information provided by large language models affects human decision-making processes.

Consistency of behavior patterns:

It is expected that under different simulated financial scenarios, participants' investment behaviors will show some consistency, reflecting the universality of cognitive bias in human decision-making.

Qualitative analysis of expected results

Expression of cognitive bias:

In the content analysis of human-computer dialogue, it is expected that significant expression of cognitive bias can be identified. For example, when the model provides information that conflicts with participants' preconceived views, they may exhibit confirmation bias, ignoring or questioning information that doesn't match their expectations.

Trust in large language models:

Participants are expected to express varying degrees of trust in the accuracy and reliability of large language models during the experiment. This degree of confidence may vary depending on the consistency of the information provided by the model with the actual investment results.

Complexity of decision reasoning:

Through topic coding, different levels of complexity are expected to be found in participants' decision reasoning, including intuitive responses, logical analysis, and a comprehensive consideration of model recommendations.

Overall, this research is expected to shed light on the dynamics of cognitive bias in AI-human interactions and provide empirical support for theoretical models in behavioral finance. We also anticipate that these findings will contribute to the understanding of how large language models influence financial decision-making processes, potentially revealing new human-machine alignment strategies to reduce the negative impact of cognitive bias on financial decision-making."

If the results are as expected, this could have far-reaching implications for financial professionals, investor education, and fintech development. For example, financial services providers may need to design better user interfaces to reduce the impact of cognitive biases, while investor education programs should include how to identify and respond to these biases.

Research Timeline:

September 2024 (preparatory phase)

Identify research teams and role assignments

Design experiments and research methods

Develop and test the experimental platform

Prepare and submit ethics review requests

October (Start-up phase)

Obtain ethical review approval

Start recruiting participants

Conduct initial training and pilot studies

The experimental design was optimized according to the experimental results

November (Data Collection Phase I)

Data collection is started

Monitor experiment progress to ensure data quality

Process preliminary data and feedback in real time

December (Data Collection Phase II)

Continue to collect data

Start preliminary data analysis

Prepare year-end progress reports

January 2025 (Data analysis Phase I)

Complete all data collection

Perform detailed data cleaning and sorting

Start quantitative and qualitative data analysis

February (Data Analysis Phase II)

In-depth data analysis

Identify preliminary research findings

Prepare an interim study report

March (Writing and review phase)

Start writing a research paper

Organize regular team meetings to discuss research progress

Submit a draft paper for internal review

April (Revision and submission phase)

Revise the paper based on feedback

Prepare the final research report and presentation materials

Submit papers to journals/conferences

Arrange project review meetings and experience sharing

Practical progress is subject to change due to unforeseen circumstances and adjustments encountered during the study. Given the time taken for review and publication in academic journals, actual progress may be extended or shortened.

Conclusion:

This research program aims to explore how cognitive bias affects financial decision-making and further examine the role of artificial intelligence, in particular large language models, in this process. Through carefully designed experimental methods and detailed data collection, this study is expected to reveal a series of key findings that are not only critical to the theoretical development of behavioral finance, but also have profound implications for applications in the field of artificial intelligence.

Possible key findings:

The relationship between cognitive bias and investment behavior: We expect to find out how specific cognitive bias is related to individual investment behavior patterns, so as to provide empirical support for understanding investors' irrational behavior.

Ai's role in behavior modification: Possible findings also include how large language models influence and potentially correct for these cognitive biases, revealing AI's potential in promoting more rational financial decision-making.

Contributions to behavioral finance:

This research will enrich the literature of behavioral finance, particularly in understanding the process of how individuals make decisions in complex financial environments. By combining traditional behavioral finance theory with the latest artificial intelligence technology, we are able to provide a more comprehensive perspective to observe and analyze investor behavior.

Contributions to artificial intelligence:

This study will also demonstrate the effectiveness of large language models in processing complex decision tasks, providing valuable insights for developing more efficient AI-assisted decision systems. This will help promote the application and development of artificial intelligence technology in the financial field. It also provides a new approach to alignment based on human behavioral finance.

Practical significance of the study:

Our research has immediate practical implications for financial professionals, policy makers, and ordinary investors. The results can help design better investment strategies and improve the efficiency of financial markets, as well as serve as a scientific basis for financial education and investor protection policies.

Implications for future research:

This research program will also provide new directions for future research in the field of behavioral finance and artificial intelligence. For example, future research could further explore the role of different types of cognitive bias in other financial decision situations, or assess the effectiveness of other types of AI techniques in assisting financial decision making.

In summary, this research program will help advance our understanding of the intersection of behavioral finance and artificial intelligence, and its findings are expected to provide valuable theoretical and practical insights and lay a solid foundation for future research efforts.

Reference:

[1] Cao, L. (2022). Ai in finance: challenges, techniques, and opportunities. ACM Computing Surveys (CSUR), 55(3), 1-38.

[2] Li, Y., Wang, S., Ding, H., & Chen, H. (2023, November). Large Language Models in Finance: A Survey. In Proceedings of the Fourth ACM International Conference on AI in Finance (pp. 374-382).

[3] Hadi, M. U., Qureshi, R., Shah, A., Irfan, M., Zafar, A., Shaikh, M. B., ... & Mirjalili, S. (2023). A survey on large language models: Applications, challenges, limitations, and practical usage. Authorea Preprints.

[4] Wu, S., Irsoy, O., Lu, S., Dabravolski, V., Dredze, M., Gehrmann, S., ... & Mann, G. (2023). Bloomberggpt: A large language model for finance. arXiv preprint arXiv:2303.17564.

[5] Svetlova, E. (2022). AI ethics and systemic risks in finance. AI and Ethics, 2(4), 713-725.

[6] Anshari, M., Almunawar, M. N., Masri, M., & Hrdy, M. (2021). Financial technology with ai-enabled and ethical challenges. Society, 58(3), 189-195.

[7] Johnson, K., Pasquale, F., & Chapman, J. (2019). Artificial intelligence, machine learning, and bias in finance: toward responsible innovation. Fordham L. Rev., 88,

499.

[8] Lin, T. C. (2019). Artificial intelligence, finance, and the law. Fordham L. Rev., 88, 531.

[9] Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases: Biases in judgments reveal some heuristics of thinking under uncertainty. science, 185(4157), 1124-1131.

[10] Zhang, Y., & Zheng, X. (2015). A study of the investment behavior based on behavioral finance. European Journal of Business and Economics, 10(1).

[11] Ahmad, Z., Ibrahim, H., & Tuyon, J. (2017). Institutional investor behavioral biases: syntheses of theory and evidence. Management Research Review, 40(5), 578-603.

[12] Liu, H., Peng, C., Xiong, W. A., & Xiong, W. (2022). Taming the bias zoo. Journal of Financial Economics, 143(2), 716-741.

[13] Hagendorff, T., Fabi, S., & Kosinski, M. (2023). Human-like intuitive behavior and reasoning biases emerged in large language models but disappeared in ChatGPT. Nature Computational Science, 3(10), 833-838.

[14] Su, J., Lang, Y., & Chen, K. Y. (2023). Can AI Solve Newsvendor Problem Without Making Biased Decisions? A Behavioral Experimental Study. A Behavioral Experimental Study (September 1, 2023).

[15] Jones, E., & Steinhardt, J. (2022). Capturing failures of large language models via human cognitive biases. Advances in Neural Information Processing Systems, 35, 11785-11799.

[16] Ferrara, E. (2023). Should chatgpt be biased? challenges and risks of bias in large language models. arXiv preprint arXiv:2304.03738.

[17] Schramowski, P., Turan, C., Andersen, N., Rothkopf, C. A., & Kersting, K. (2022). Large pre-trained language models contain human-like biases of what is right and wrong to do. Nature Machine Intelligence, 4(3), 258-268.

[18] Kirk, H. R., Jun, Y., Volpin, F., Iqbal, H., Benussi, E., Dreyer, F., ... & Asano, Y. (2021). Bias out-of-the-box: An empirical analysis of intersectional occupational biases in popular generative language models. Advances in neural information processing systems, 34, 2611-2624.

[19] Wang, N., Yang, H., & Wang, C. D. (2023). FinGPT: Instruction Tuning Benchmark for Open-Source Large Language Models in Financial Datasets. arXiv preprint arXiv:2310.04793.

[20] Wang, Y., Zhong, W., Li, L., Mi, F., Zeng, X., Huang, W., ... & Liu, Q. (2023). Aligning large language models with human: A survey. arXiv preprint arXiv:2307.12966.

[21] Rahwan, I., Cebrian, M., … Wellman, M. (2019). Machine behaviour. Nature, 568(7753), 477–486. https://doi.org/10.1038/s41586-019-1138-y

[22] Tinbergen, N. (2005). On aims and methods of Ethology. Animal Biology, 55(4), 297–321. https://doi.org/10.1163/157075605774840941

[23] Yudkowsky, E. (2016). The AI alignment problem: why it is hard, and where to start. Symbolic Systems Distinguished Speaker, 4.

[24] Christian, B. (2020). The alignment problem: Machine learning and human values.

WW Norton & Company.

[25] Kahneman, D., & Tversky, A. (2013). Prospect theory: An analysis of decision under risk. In Handbook of the fundamentals of financial decision making: Part I (pp. 99-127).

[26] Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. Cognitive psychology, 3(3), 430-454.

[27] Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. Cognitive psychology, 5(2), 207-232.

[28] Hon, T. Y., Moslehpour, M., & Woo, K. Y. (2021). Review on behavioral finance with empirical evidence. Advances in Decision Sciences, 25(4), 1-30.

[29] Kliger, D., van den Assem, M. J., & Zwinkels, R. C. (2014). Empirical behavioral finance. Journal of Economic Behavior & Organization, 107, 421-427.

[30] Chaffai, M., & Medhioub, I. (2014). Behavioral finance: An empirical study of the Tunisian stock market. International Journal of Economics and Financial Issues, 4(3), 527-538.

[31] Alsabban, S., & Alarfaj, O. (2020). An empirical analysis of behavioral finance in the Saudi stock market: Evidence of overconfidence behavior. International Journal of Economics and Financial Issues, 10(1), 73.

[32] Zhang, B., Yang, H., Zhou, T., Ali Babar, M., & Liu, X. Y. (2023, November). Enhancing financial sentiment analysis via retrieval augmented large language models. In Proceedings of the Fourth ACM International Conference on AI in Finance (pp. 349-356).

[33] Hadfield-Menell, D., Russell, S. J., Abbeel, P., & Dragan, A. (2016). Cooperative inverse reinforcement learning. Advances in neural information processing systems, 29.

[34] Rahwan, I., Cebrian, M., Obradovich, N., Bongard, J., Bonnefon, J. F., Breazeal, C., ... & Wellman, M. (2019). Machine behaviour. Nature, 568(7753), 477-486.

[35] Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. The journal of Finance, 25(2), 383-417.

[36] Barberis, N., & Thaler, R. (2003). A survey of behavioral finance. Handbook of the Economics of Finance, 1, 1053-1128.

[37] Moore, D. A., & Healy, P. J. (2008). The trouble with overconfidence. Psychological review, 115(2), 502.

[38] Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. Review of general psychology, 2(2), 175-220.

[39] Shefrin, H., & Statman, M. (1985). The disposition to sell winners too early and ride losers too long: Theory and evidence. The Journal of finance, 40(3), 777-790.

[40] Rabin, M., & Schrag, J. L. (1999). First impressions matter: A model of confirmatory bias. The quarterly journal of economics, 114(1), 37-82.

[41] Lerner, J. S., Li, Y., Valdesolo, P., & Kassam, K. S. (2015). Emotion and decision making. Annual review of psychology, 66, 799-823.

[42] Malmendier, U., & Tate, G. (2005). CEO overconfidence and corporate investment. The journal of finance, 60(6), 2661-2700.

[43] Baker, M., Ruback, R. S., & Wurgler, J. (2007). Behavioral corporate finance. In

Handbook of empirical corporate finance (pp. 145-186). Elsevier.

[44] Bikhchandani, S., Hirshleifer, D., & Welch, I. (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. Journal of political Economy, 100(5), 992-1026.

[45] Camerer, C. F. (2003). Behavioural studies of strategic thinking in games. Trends in cognitive sciences, 7(5), 225-231.

[46] Lusardi, A., & Mitchell, O. S. (2014). The economic importance of financial literacy: Theory and evidence. American Economic Journal: Journal of Economic Literature, 52(1), 5-44.

[47] Lusardi, A., Mitchell, O. S., & Curto, V. (2014). Financial literacy and financial sophistication in the older population. Journal of pension economics & finance, 13(4), 347-366.

[48] Loughran, T., & McDonald, B. (2016). Textual analysis in accounting and finance: A survey. Journal of Accounting Research, 54(4), 1187-1230.

[49] Buchanan, D. A., & Huczynski, A. (2019). Organizational behaviour. Pearson UK.

[50] Zhang, B. H., Lemoine, B., & Mitchell, M. (2018, December). Mitigating unwanted biases with adversarial learning. In Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society (pp. 335-340).