

Artificial Intelligence and Behavioural Finance: Correcting Biases or Creating New Risks?

Vihaan Jain

Student, Neeja Modi School

Abstract

The rapid emergence and integration of Artificial Intelligence (AI) in the financial sector have led to many concerns regarding the potential impact on investors' behavior and decision-making. This paper attempts to resolve the research problem related to the role and potential impact of AI in the field of behavioral finance. The research focuses on the following aspects through a case study approach: AI-based robo-advisors in Southeast Asia, sentiment analysis through the “photo pessimism” study, and ethical concerns in generative AI-based financial advice. The results show that AI can be an effective tool in removing emotional biases from investors' behavior, such as panic selling and overconfidence, by adopting a disciplined approach to investments. Moreover, it helps in better understanding market sentiments through the analysis of data from photographs and media, indicating the role of group emotions in determining market sentiments. However, there are many challenges and concerns related to the role and potential impact of AI in the field of behavioral finance. The findings suggest that AI can play a significant role in improving efficiency, enhancing financial decision-making, mitigating investor biases, and interpreting market sentiment, while also introducing ethical and transparency challenges.

Keywords: Artificial Intelligence, Behavioural Finance, Robo-Advisors, Sentiment Analysis, Explainable AI

1. Introduction

Artificial Intelligence (AI) has evolved drastically over the past few years, rebuilding industries across the world, including the financial sector. Originally, financial decision-making was based upon human judgment, human biases and fundamental or technical analysis. People used to make decisions by considering the available information about the financial market and the trades they had to make. However, the involvement of AI in everyone's day-to-day life has led to a change in the investors' operations, as they have started to switch to AI-based decision-making more than human-based judgments (Raso et al., 2018). This is because these AI systems process large amounts of data very quickly, allowing for faster and quicker financial decisions. AI has been involved in the financial sector in various ways, such as through systematic trading, portfolio management, fraud prevention, and automatic advisory services. Financial institutions across countries have used AI technologies to improve productivity and reduce costs. As a result, AI is no longer considered an untrusted tool, but rather a significant part of financial systems (Turksen et al., 2024). Nowadays behavioral finance is also being impacted by the intervention of AI and being taken into more consideration, as it explains why investors often tend to behave irrationally during periods of uncertainty, market bubbles, and financial crises. Investors tend to adhere to irrational behaviours such as herd mentality, overconfidence, loss aversion, and panic selling,

especially during market fluctuations (Umeaduma, 2024). However, now researchers across various countries have started to incorporate the psychological aspects into financial models, making behavioural finance an increasingly relevant and expanding discipline (Muradoglu & Harvey, 2012).

Behavioral finance and AI together have introduced new methods for financial analysis to analyze investor sentiment, emotions, and decision patterns. Through the use of machine learning, and deep learning techniques AI systems can now analyze investor sentiment through language processing, interpreting visual information and identifying behavioural patterns. These advancements allow emotional and behavioural biases to be measured easily as earlier they weren't even considered or measured. Yet, the high usage of AI in the financial world has increased several concerns amongst people about transparency, data privacy, and overreliance on AI systems. Many AI models function as “black boxes,” meaning users cannot clearly understand why certain financial recommendations are being given, which creates issues of accountability and trust (Hassija et al., 2024). Thus, as AI increases efficiency in the financial world, it also brings along new risks and ethical challenges. This paper explores how AI and behavioural finance work hand in hand in shaping the user's financial behaviour while also bringing out challenges that have to be considered.

This paper examines the relationship between AI and behavioural finance by exploring three themes, each with case studies. The first theme showcases how AI helps in detecting and correcting human (investor) biases, using the robo-advisory platforms in Southeast Asia. This case sheds light on how these systems help to reduce emotional biases by motivating one to adhere to long-term investment plans and other strategies regardless of the fluctuations in the market. This basically helps the investors with limited financial knowledge to get exposure to this new spectrum. The second theme focuses on AI's role in interpreting market behaviour and behavioural patterns through visual and media-based data. By analyzing the news images and emotional sentiments, AI systems hypothesize the market movement and predict the way the investor psychology as a whole will work. This theme is aided by “photo-pessimism,” which portrays how the negative picture leads to a predictable downtrend in the market returns. The third theme brings out the main challenges and risks due to using AI in behavioural finance. This theme highlights ethical issues such as a lack of transparency, privacy risk, and the potential for misuse of data. Thus, these themes display a neutral understanding of AI's intervention in behavioural finance, improving it, along with the comprehensive view of how the growing role of AI in today's world plays a role in shaping financial behaviour.

2. AI Detecting and Correcting Investor Biases

Over the past few years, financial markets have changed dramatically. Investors in South Asia deal with a lot of financial problems, mostly because they are not exposed to the financial market. Many amateur investors enter the market without strong financial literacy, relying heavily on social media trends or peer influence (Coumans, 2025). Adding on, emotions tend to alter thoughts a lot, like when, during periods of market trepidation, investors just sell off everything during a market dip, or overconfidence makes them jump on whatever trend looks favourable. This behaviour is generally seen during market crashes or bullish trends, where emotional reactions take over logical reasoning. Thus, behavioural finance explains why these biases affect the decisions and prevent investors from using their knowledge properly. Theories such as herd behaviour, loss aversion, and anchoring bias explain why investors fail to act prudently even when the required information is available to them.

Studies suggest that robo-advisors are powered by AI (Aini, 2025). These robo-advisors adhere to algorithmic portfolio building and automatic asset allocation techniques. AI looks at things through the information they have about the users' investing goals, their risk tolerance, and how long they can hold onto goods. It processes this information using machine learning to match users with low-risk investment portfolios that help them accomplish their financial objectives. Then, the AI puts together the portfolio and sticks to it, not letting some minor changes affect its choices. This systematic approach reduces the impact of short-term volatility. Unlike humans, who tend to panic and let emotions take over by acting in those moments, the system adheres to its roots.

In Southeast Asian countries like Singapore, there are tons of new investors, unaware of the financial market, who may blindly trust the society and be unaware of the risks of losing money, so AI could prevent the unstable choices (Coombs & Redman, 2018). It encourages staying with the plan for a longer period, which is very difficult for individuals to follow. Nevertheless, the article points out problems of AI being a “black box” where users cannot figure out exactly why it chooses one investment over another. That lack of clarity might prohibit one from trusting the ‘robo-advisor’. Moreover, not everyone gets access to these advanced tools, usually just the wealthier people, and that could widen the gap between different groups even more. This creates a potential inequality in financial advantage, where technologically equipped investors benefit more than others, increasing socio-economic disparities.

Overall, AI does help cut down on emotional biases, improving the investment decision. It introduces structure and discipline into personal finance. However, combining it with human intelligence is a safer choice, allowing investors to still be involved and not just turn everything over blindly. A hybrid approach, where AI provides analytical support while humans maintain strategic monitoring, will create a more balanced and worthy investment.

3. AI Interpreting Market Sentiment and Behavioural Patterns

In a fast-paced modern world, AI has been contributing to behavioural finance. With the explosion of digital media, financial markets are now heavily influenced by real-time information flow, making sentiment analysis increasingly important. This is done by sentiment analysis. Sentiment analysis refers to the use of computational models to identify and measure emotions, tone, and attitudes within large datasets such as news articles, social media posts, and visual content (Rout et al., 2018). For example, a photo pessimism study reveals how AI can understand the emotional signals of humans through the news images rather than relying upon text messages. This portrays a shift from old text-based sentiment analysis to more advanced visual emotion recognition systems. Investors tend to react in a different way: news can trigger various emotions like fear, happiness, overconfidence, etc. These emotional reactions often occur subconsciously, influencing trading decisions even before one's cognitive processing of information. This study involves using CNN to analyze multiple photos altogether, which allows researchers to form an index – “photo pessimism” – to predict the market's turns. Convolutional Neural Networks (CNNs), a deep learning technique commonly used in image recognition, are trained to detect patterns such as facial expressions, environmental context, and visual tone within news images. By using all of these patterns, researchers made an emotional intensity index that captures the overall pessimism reflected in media coverage, such as news, etc.

Griffith et al. (2020) show how the stock market returns and the news images share a direct relation, where when the news is more negative, the market returns tend to fall by a measurable amount, which shows how the investors acted emotionally. The findings indicate that financial markets are not fully driven by

fundamental financial indicators but are also dependent upon mood and perception. Yet, this behaviour tends to support the behavioural finance theories, which claim that emotional sentiments can distort prices. Especially, as the study suggests, the smaller stocks are more significantly impacted as they are volatile in nature too. Smaller or less liquid stocks are generally more vulnerable to sentiment shifts because they lack strong institutional backing, making them more vulnerable to emotional trading patterns.

Huang & Rust (2021) highlight how AI uses emotional information extracted from visual information that can be used as indicators of the market's position the next day. It demonstrates that AI can transform emotional cues into financial signals, thereby bridging psychology and data science. Although it brings to the table how AI has no justification behind its choice of action, and why it concludes that the investors will react in that certain way only. This raises concerns regarding deep learning systems as they often function as black-box models, making it difficult to fully understand the reasoning behind their conclusions and suggestions.

Overall, Sahani (2024) showcases that AI is not just predicting markets but revealing how human emotions, shaped by media and visuals, influence financial decision-making at a collective level. It explores that financial markets operate within a social and psychological ecosystem, where perception and emotions can be as impactful as economic fundamentals.

4. Risks and Ethical Challenges of AI in Behavioural Finance

While AI offers multiple advantages, it comes with its own drawbacks, which are highlighted throughout the article by Choukhmane, de Silva, Lin and Akuzawa (2025) “How Good is Generative AI Personal Financial Advice?” This article specifically evaluates whether generative AI systems are truly capable of delivering trustworthy and impartial financial advice, especially in multiplex and emotionally driven market environments. The most significant concern brought out through this article is bias amplification. Bias amplification refers to the process by which AI systems unintentionally reinforce and magnify existing behavioural patterns present in historical data. This happens when AI models use the previous transaction methods and adhere to them, which eventually leads to similar errors made in the past by the user due to herding, overconfidence, etc. Since machine learning models are modelled on past datasets, they may treat irrational behaviour as a normal pattern rather than identifying it as a mistake. Instead of correcting the previous errors, AI tends to adhere to them.

Another significant issue portrayed is the lack of transparency, known as the “black box” problem. Generative AI models, especially those based on deep architectures, often operate through highly complex networks that are not easily understood by users. This article conveys how the AI usually provides advice and suggestions without justifying its reason to choose that method. This is quite troublesome in the field of finance, as the author notes say “users may receive confident recommendations without understanding the assumptions behind them.” Such a lack of transparency weakens accountability and prevents investors from critically evaluating the confirmation of the advice provided. To address this issue, recent developments in Explainable Artificial Intelligence (XAI) offer a potential solution. A possible solution to the “black box” problem is the development of Explainable Artificial Intelligence (XAI), which focuses on making AI systems more transparent and understandable. Unlike traditional AI models that provide outputs without clear reasoning, XAI aims to explain why a particular financial recommendation is made. This is especially important in behavioural finance, where understanding the reasoning behind decisions is crucial. The improved interpretability can enhance user trust and reduce blind reliance on AI systems (Vijay Hassija et al. 2024). In this context, XAI can help investors evaluate whether decisions are based

on logical financial indicators or biased historical data, thereby reducing risks such as bias amplification and the principal-agent problem. However, XAI is still developing and may involve trade-offs between accuracy and interpretability, meaning further research is needed before it can be fully relied upon in financial decision-making. This leads to blind trust and overconfidence in the AI systems, leading to some severe losses, too. In high-risk financial decisions, unquestioned trust in algorithmic outputs can maximize financial risk rather than reduce it. Also, the article brings out an over-reliance on past behaviour, which suggests how one's mindset may alter with time, due to which it may have a psychological development, which may alter their choices over time during crises or boom periods. Human tastes, risk tolerance, and emotional biases are dynamic and can shift significantly depending on life circumstances or macroeconomic events. Moreover, during such market volatile conditions, AI tends to fail as they adhere to past patterns, so it can't predict the future market changes at such moments better than human technical and fundamental analysts. Extreme market disturbance, such as financial crises or unplanned geopolitical shocks, often breaks historical trends, limiting the power of data-dependent AI systems.

The literature also discusses principal-agent problems in the realm of AI, where the AI knows more than the investor, and it may force the investor to act in a way that benefits the market or that specific company more than the investor. This asymmetry of information can create conflicts of interest, especially if AI systems are developed or sponsored by financial institutions with their own incentives, leading to losses to investors who blindly trust AI models. Moving further, another concern brought up is about data privacy. Because the AI systems demand sensitive financial data from the user, this may lead to further risks of data misuse. Personal financial data, spending attitude, and risk tolerance data are highly secretive, and any misuse can have serious problems for investors. Lastly, through the user and server interactions, AI may also interpret the fear tolerance level of the user, which may lead to manipulation and exploitation of the user. Such a behaviour profile could potentially be used to nudge investors toward decisions that align more with platform incentives rather than the investors' incentives.

Overall, the case explored how AI effectively supports financial decision making, yet it must be used with proper surveillance, as when AI is not monitored with regular intervals, it may worsen the behavioural problems of the user. Therefore, moral rules, transparency standards, and continuous human monitoring become essential to ensure that AI remains a helpful tool rather than a way of systemic behavioural risk.

5. Conclusion

To conclude, the present study talks about how Artificial Intelligence is remodelling the field of behavioural finance by impacting the investor biases and the way they make decisions. It set out to find out whether AI improves computational efficiency in finance or whether it fundamentally changes the psychological base of investment attitude. By using real-world case studies, this paper explores the opportunities and challenges of AI getting involved in the field of behavioural finance. Through a thematic structure, the study critically explored both the advantageous and disadvantageous sides of AI integration with behavioural finance.

The 3 themes of this paper show that AI is not just a technical tool, but also an asset when it comes to behavioural finance to an extent. The results show that AI works with the combination of both technology and psychology, impacting not only data processing but also human cognition. Through the case of robo advisors in Southeast Asia, it becomes very much evident of how AI can control emotional biases like panic selling, overconfidence, and herd behaviour. These AI involved investors can adhere to their strategies of discipline and long-term thinking, mainly helping people new to the financial market. This

portrays that planned algorithmic systems can reduce irrational behaviour and encourage stability, particularly among inexperienced retail investors.

Moreover, the second theme showed how AI can interpret market sentiment using media-based information. The study of photo pessimism portrays how the images of news triggered human emotions which affect the market returns and investor behaviour strongly. This piece of evidence reinforces behavioural finance that markets are influenced not only by fundamentals but also by collective mood and perception.

However, the third theme reveals that the use of AI and behavioural finance brings in some problems and major concerns like data privacy, transparency, and overreliance, which can harm the user if not used carefully. Issues such as bias amplification, black-box decision-making, and principal-agent conflicts indicate that AI systems may introduce new structural risks if not regulated properly.

Hence, this paper is a reflection on AI's role in behavioural finance and how AI has strong potential to improve efficiency and decision-making, yet it can't replace human judgment. Instead, it should be used as a supportive tool for financial education and advice to a limited extent. The present study is primarily based on secondary case analyses and theoretical interpretation, which limits the ability to draw all possible conclusions. Future research should involve primary data sources and long-term performance analysis of AI-driven advisory systems. Therefore, the conclusions drawn in this paper should be understood as a critical foundation for further exploration rather than a definitive judgment on AI's role in finance.

References

1. Aini, S. (2025). Trusting AI in Finance: The Role of Psychological Traits and AI Literacy in Shaping User Behaviour Toward Robo-Advisors. *The South East Asian Journal of Management*, 19(2), 4.
2. Coombs, C., & Redman, A. (2018). The impact of robo-advice on financial advisers: A qualitative case study.
3. Coumans, J. (2025). Beyond Fundamentals: The Role of Peer Influence, Salience, and Financial Literacy on Meme Stock Investing.
4. Griffith, J., Najand, M., & Shen, J. (2020). Emotions in the stock market. *Journal of Behavioral Finance*, 21(1), 42-56.
5. Hassija, V., Chamola, V., Mahapatra, A., Singal, A., Goel, D., Huang, K., ... & Hussain, A. (2024). Interpreting black-box models: a review on explainable artificial intelligence. *Cognitive Computation*, 16(1), 45-74.
6. Huang, M. H., & Rust, R. T. (2021). A strategic framework for artificial intelligence in marketing. *Journal of the academy of marketing science*, 49(1), 30-50.
7. Muradoglu, G., & Harvey, N. (2012). Behavioural finance: the role of psychological factors in financial decisions. *Review of Behavioural Finance*, 4(2), 68-80.
8. Raso, F. A., Hilligoss, H., Krishnamurthy, V., Bavitz, C., & Kim, L. (2018). Artificial intelligence & human rights: Opportunities & risks. Berkman Klein Center Research Publication, (2018-6).
9. Rout, J. K., Choo, K. K. R., Dash, A. K., Bakshi, S., Jena, S. K., & Williams, K. L. (2018). A model for sentiment and emotion analysis of unstructured social media text. *Electronic Commerce Research*, 18(1), 181-199.
10. Sahani, T. (2024). Decoding market emotions: the synergy of sentiment analysis and AI in stock market predictions. *Journal of Next-Generation Research* 5.0.

11. Turksen, U., Benson, V., & Adamyk, B. (2024). Legal implications of automated suspicious transaction monitoring: enhancing integrity of AI. *Journal of Banking Regulation*, 25(4), 359-377.
12. Umeaduma, C. M. G. (2024). Behavioral biases influencing individual investment decisions within volatile financial markets and economic cycles. *Int J Eng Technol Res Manag*, 8(03), 191.