IJFBS
Finance & Banking Studies

Finance & Banking Studies

IJFBS VOL 14 NO 1 ISSN: 2147-4486

Available online at www.ssbfnet.com
Journal homepage: https://www.ssbfnet.com/ojs/index.php/ijfbs

Bridging Human Expertise and AI: Evaluating the Role of Large Language Models in Retail Investors' Decision-Making

Dmitrii Gimmelberg [(a,b)], Marta Głowacka [(b)], Alexey Belinskiy [(b,c)], Sergei Korotkii [(b)], Valentin Artamov [(b)], Iveta Ludviga [(a)]*

[(a)] *RISEBA University of Applied Sciences, Faculty of Business and Economics Latvia.*
[(b)] *Aestima SIA, Research and Development, Latvia.*
[(c)] *SBS Swiss Business School, Doctorate Studies, Switzerland.*

A R T I C L E I N F O

Article history:

Received 23 Dec. 2024
Received in rev. form 03 Feb. 2025
Accepted 03 Feb. 2025

Keywords:

LLM, Large Language Models, Retail Investors, Investments

JEL Classification:
G41, C81

A B S T R A C T

This study investigates the role of large language models (LLMs) in retail investors (RIs) decision-making processes from the perspective of the Theory of Planned Behaviour (TPB). It explores whether LLMs can replace or change the role of financial experts and whether introducing LLMs may lead to more infromed RIs' decisions. Qualitative interviews were conducted with experienced RIs (n = 8). Secondary data were gathered from YouTube recordings (n = 44). Thematic analysis and Retrieval-Augmented Generation (RAG) methodology was used for data extraction and analysis of the scripts. The findings indicate that while LLMs have the potential to enhance accessibility to expert opinions and provide more informed investment decisions, they are unlikely to replace human experts. RIs show a preference for combining LLM insights with human expertise, recognising the limitations of LLMs in managing complex and nuanced investment information. The study highlights the usefulness of the TPB as a framework for the exploration of the topic. It also introduces a novel research method - advanced data extraction techniques on a vast unstructured dataset. Results contribute to the understanding of LLMs potential in supporting RIs and confirms the usefulness on the AESTIMA tool for data extraction and analysis.

Introduction

Retail and individual investors' decisions to buy or sell financial market instruments are influenced by opinions from a range of sources, including information from financial experts, brokers, big investors (Bikhchandani & Sharma, 2000; Islamoğlu et al., 2015; Patil & Bagodi, 2021), friends (Dierkes et al., 2009) and family (Mahalakshmi & Anuradha, 2018), and expert opinions disseminated through media, including social media (Patil & Bagodi, 2021; Subramanian, 2021). Some investors, however, prefer to base their decisions on financial and fundamental analysis (Shefrin & Statman, 2000). Reliance on experts' opinions can create biases such as anchoring, where investors adjust their portfolios based on those opinions rather than an independent analysis (Campbell & Sharpe, 2009), especially during times of market optimism. Biased, sensationalised news or viral social media posts can trigger emotional reactions and cause stock price volatility, leading to emotional decisions rather than in-depth analysis of data, resulting in impulsive investment choices (Daudert, 2021). The fake or poorly verified news can lead to market instability, investors' losses and to prices of assets reaching a different value to their true value (Sari et al., 2022).

LLMs can be transformative tools in the finance sector, offering a range of applications that enhance efficiency, accuracy, and decision-making processes (X. Zhao et al., 2023). LLMs can extract detailed insights from financial documents, conduct sentiment analysis, forecast market trends, provide a comprehensive assessment of financial health and market conditions (Cao & Feinstein, 2024; Li et al., 2023; Ma et al., 2024) to help investors and financial institutions make informed decisions (Guo et al., 2023). LLMs can also assist in interpreting complex regulatory documents, supporting compliance with legal requirements (Guo et al., 2023).

* *Corresponding author.*
https://doi.org/10.20525/ijfbs.v14i1.3910

LLMs can support risk assessment and fraud detection due to their ability to identify patterns and anomalies in large sets of financial data, highlighting possible fraudulent activities (Raiaan, et al., 2024), which is especially important in loan approval or money laundering detection (Chen et al., 2024; Lo & Ross, 2024). LLMs can optimise investment strategies by analysing market data and generating actionable insights (Li et al., 2023; Lo & Ross, 2024). Their capability to understand natural language instructions may be particularly useful in customer service (Raiaan, et al., 2024).

Despite the potential of LLMs, there is still a lack of knowledge on their role in investment decisions. Investors need to critically evaluate the sources of information and consider a balanced approach that includes both external advice and independent analysis. The potential of LLMs is still evolving, and in future, they might replace humans in some areas of professional practice. Understanding the motivations behind the introduction or resistance to introducing this technology into RIs practice is sparse. Gaining better insight into the issue would guide the development of LLMs based tools for the industry and formulate strategies for their implementation. This research aims to investigate the role of LLMs in the decision-making process of retail investors (RIs) through answering two research questions: 1) Are LLMs seen as a potential replacement of human financial experts, or rather will the experts' role evolve alongside LLMs introduction to the industry? 2) Does LLM use improve decision-making by enabling more informed, rational investments?

The second objective of this research is to validate the quality of data extraction and analysis from unstructured text sources, such as YT recordings scripts, using the Retrieval-Augmented Generation (RAG) (Gao et al., 2024; Yu, 2022); using AESTIMA tool[1]. This objective supports the primary research by ensuring that the data used to analyse the impact of LLMs on RIs behaviour was robust.

Theoretical Framework

The Theory of Planned Behaviour (TBP) (Ajzen, 2011; Schifter & Ajzen, 1985) has been employed in previous studies in the field of finances for example to understand casual investors' investment decisions (Sharma & Gupta, 2011), to explore the impact of psychological factors on capital market investments (Piyumini & Wijethunga, 2020), predict people's intentions to invest in cryptocurrencies (Norisnita & Indriati, 2022) understand mutual funds' investments (Dewi & Ronny, 2023), or to predict the intention to trade online (Gopi & Ramayah, 2007). A systematic literature review applied the TPB to develop a theory of financial planning behaviour (Yeo et al., 2023). The TPB was applied to investors' decisions in studies conducted in different settings such as stock market (Pahlevi & Oktaviani, 2018; Pham et al., 2021), or mutual funds (Widyastuti et al., 2023). The studies included populations, such as Millennials (Setyorini & Indriasari, 2020), Gen Z and Gen Y (Gumasing & Niro, 2023; Malzara et al., 2023; Nugraha & Rahadi, 2021; Nugraha & Prasetyaningtyas, 2023) professionals from investment sector (Adam & Shauki, 2014), or private households (Gamel et al., 2022).

According to the TPB, intention to perform behaviour is affected by three factors: attitude, subjective norm (SN) and perceived behavioural control (PBC). Attitude is a tendency towards an object, person, or idea, which has cognitive, and affective components and contributes to a person's behaviour (Ajzen, 2011; Schifter & Ajzen, 1985). Attitudes towards saving, borrowing, or investing were reported to influence intention to engage in those behaviour (Kisaka, 2014; Pham et al., 2021). Investors' attitudes towards risk, return, and other financial factors also influence their investment decisions (Rooh et al., 2023). SN covers social rules and pressures from an individual social environment (Ajzen, 2011; Schifter & Ajzen, 1985), which can influence investment intentions (Nugraha & Prasetyaningtyas, 2023; Pahlevi & Oktaviani, 2018; Sharma & Gupta, 2011). On the other hand, a study exploring investment decisions on the crowdfunding platform reported that SN did not affect behaviour (Yulandreano & Rita, 2023). SN can also guide investors towards ethical financial decisions (Adam & Shauki, 2014). PBC refers to an individual's perception of how easy or difficult performing certain behaviour can be (Ajzen, 2011; Schifter & Ajzen, 1985). Multiple studies have shown that PBC influences investment intentions and behaviours, for example, it positively impacted the crowdfunding intentions (Yulandreano & Rita, 2023) or green investment intentions (Malzara et al., 2023).

Research Methodology

The qualitative exploration of the topic allows gaining insights supporting in depth understanding of the topic with consideration for the context and nuance in the data (Braun & Clarke, 2022; Clarke & Braun, 2013). As LLMs are still a novel tool, especially for individual RIs, this exploratory approach is best fitted to answer the research questions.

Data

Two complementary sources of information were used to answer the research questions. The primary source was semi-structured interviews with experienced industry members. The interview schedule included questions formulated in line with the TPB framework and included up to 20 questions[2] aimed to capture interviewees' attitudes, SN, PBC and explore their intention to use LLMs in professional practice. The interviewees were also asked about their actual experience using LLMs professionally. Participants were eight experienced industry members (all of them were RIs, 5 were also financial professionals). They were recruited using convenience approach through the personal network of a lead researcher. All eight were male and their age range was 30 – 55 years of age. Two interviews were conducted in person and six online. Consent was obtained from all participants.

[1] https://aestima.io

[2] Some questions were found to bring redundant information and were removed from subsequent interviews.

The second data source was YouTube (YT). It was searched for videos with interviews or talks, discussing the use of LLMs within the financial markets and investment sector. 78 relevant channels, discussing the capital and financial market issues, were selected. This selection ensured representation of the mainstream financial media (Bloomberg, Yahoo Finance, CNBC etc) as well as smaller organisations, independent experts, and financial journalists. The total number of 4513 sources (transcripts) uploaded to YT in the 2nd quarter of 2024 and 1617 sources uploaded in the 2nd quarter of 2023 were gathered to be checked for their relevance to answer the research questions.

Data analysis

Data was analysed in two stages; first automated stage was completed using the AESTIMA tool followed by the manual check of the accuracy of the analysis. In accordance with RAG process operationalised through AESTIMA tool[3], following embedding and similarity search, the sources were checked for relevance to answer the research questions (Gimmelberg et al., n.d.).

A RAG approach was employed to extract and analyse data from a collection of video transcripts (Gao et al., 2024; Yu, 2022; Zhao et al., 2024). It was implemented in two stages. First, textual content (transcripts for videos) was converted into dense embeddings using the 'text-embedding-ada-002' model[4]. This type of model creates high-dimensional vector representations of text, encapsulating semantic meaning, which allows for efficient transformation of the text into embeddings (H. Cao, 2024; Keraghel et al., 2024; Petukhova et al., 2024). Next, a similarity search was conducted using Pinecone.io's automated cosine similarity algorithm[5]. The second stage involved applying the user's query and customised ranking process to the text chunks identified as most relevant in the previous step. For this task, the ChatGPT-4o-mini and ChatGPT-4o models [6] were used. This model was used to extract and analyse information from the selected text chunks, utilising its advanced language understanding and generation. By interpreting and synthesising the retrieved data, the model provided comprehensive and contextually appropriate responses to the user's queries.

The steps of relevance check are presented in Figure 1. Queries were applied to the 4500 tokens of each text, which is about 3375 words for English texts to complete the checks. The final 10 questions reflected the questions in the interviews; however, those that did not bring any new information during the interview stage were excluded. Out of all identified sources, 17 and 27 from 2024 and 2023, respectively, answered at least one of 10 questions or provided useful inferential information to the subject. Two types of responses were generated in the analysis: direct quotes from expert(s) opinions and inferential responses based on the content of the source. The quality of the responses was reviewed independently by three researchers using manual checks (comparing the output to the original source).

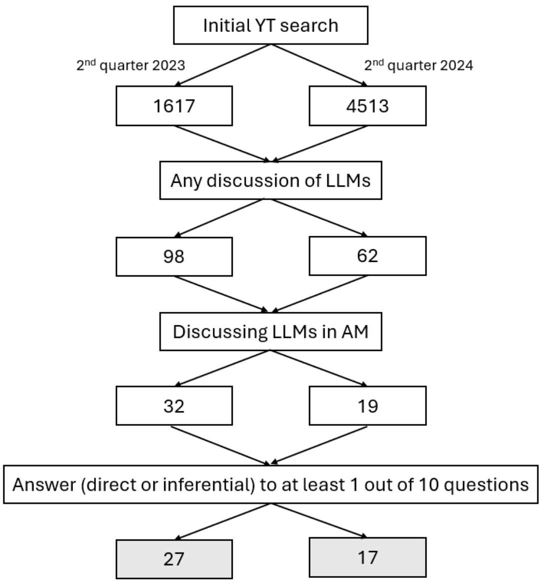

Figure 1: Relevance check flow

To evaluate the effectiveness of data extraction and analysis performed by LLMs, a random sample of 130 responses, representing 25% of the total, was selected to verify the accuracy and reliability of the automated process. The validation process was applied two criteria: (a) accuracy and thematic relevance judged by correctly understanding the subject matter, and (b) interpretive depth judged by comprehension of the extracted data's meaning. Selected statements generated by the AESTIMA tool were compared for similarity to the parts of the original text (chunks) on which these statements were based and to the entire source in manual check. The size of the chunk was fixed and was equal to 1500 tokens. One token for English texts equals approximately 0.75 of the word. The selected size of the chunk is expected to give sufficient context volume. A response was deemed accurate and confirmed if it clearly matched

[3] All prompts are available upon request.
[4] https://openai.com/index/new-and-improved-embedding-model/
[5] https://www.pinecone.io/learn/vector-similarity/
[6] https://platform.openai.com/docs/models

the original source content. It was considered partially confirmed if the meaning was preserved but some key message was omitted or not conveyed clearly. The response was marked as unconfirmed if it did not align with the original source. Any judgment discrepancies between the researchers were discussed until a consensus was reached on the final classification of the responses.

The TBP (Ajzen, 2011; Schifter & Ajzen, 1985) was selected to provide a comprehensive framework for the analysis. Both the interview responses and YouTube video transcripts were systematically analysed following a thematic analysis approach (Braun & Clarke, 2022; Clarke & Braun, 2013; Neuendorf, 2018; Vaismoradi et al., 2016); and organised into key themes and mapped onto the TPB framework. The mapping exercise provided insight into the role of expert opinions and the potential role of LLMs in RIs decisions. The synthesis involved integrating findings from both interview data and YouTube video content with a special focus on any new insights into the topic brought by the latter. The analysis was a collaborative process in which regular discussions within the research team was held to agree on themes, synthesise the data and formulate conclusions.

Results

The analysis of the interviews and YT sources revealed a range of attitudes, provided insight into SN related to the perception of roles of human experts and LLMs in RIs practice and shed light on PCB in relation to LLMs use in the investment industry.

Attitude toward behaviour

The attitude refers to the views on the possibility of replacement of human experts by LLMs in providing investment advice and the potential of LLMs to improve RIs' decisions. The general attitude towards LLMs seems to be cautiously positive. The interviewees raised issues such as LLM capabilities and limitations, democratisation of the retail investment industry, risks of widespread use of LLM to inform investment decisions, as well as issues of trust and directions of change in investment professionals' roles. In both interviews and YT recordings, introducing LLMs into practice seems inevitable although not imminent. There is recognition of LLMs' efficiency and superior data processing capabilities, but also an acknowledgement of the irreplaceable human qualities in judgment and understanding.

Many experts recognise that LLMs capabilities exceed those of human experts, especially in quick and accurate processing and analysing vast amounts of data (YT14, YT5, YT11, YTQ3). As one of the experts emphasised:

Capabilities of LLM are significantly higher than those of human experts. Specifically of specialised LLMs and/or specialised RAGs empowered by LLMs. (INT1)

LLMs can play a tri-fold role: acting as advisors by themselves, aiding in understanding expert advice and aggregating and summarising a multitude of opinions. LLMs can improve the investment decision process by optimising trading strategies, as one interviewee explaned *"LLM can critique existing strategy parameters and suggest improvements."* (INT4). LLMs can also provide an unmatched level of customisation highlighted by an expert:

All that AI is going to do is know more about you learn more about you look at patterns and you know look at successful things in the past. (YT9)

Mixed opinions were found about whether the use of LLMs would lead to more rational investment decisions. They can help dealing with human bias and contradicting information (YTQ3), offer "a second opinion" (INT5), reduce the emotional factors (YTQ5) and improve understanding of large volumes of information (YT5, YT11, YT7). However, they will not eliminate irrational human nature; *"They [IRs] will remain irrational as they are. "*(INT1). There were also concerns regarding LLMs ability to handle complex, unstructured situations, and make nuanced judgments, as one interviewee mentioned:

So far [the capabilities of LLM] are limited, just reading and aggregation of texts. They [LLM] lack nuances and interpretation and are hallucinating and there is lack of access to the most recent events. (INT5)

Similarly, in a 2023 video discussion, the author notes LLMs' limitations, particularly in recommending individual holdings or funds (YT48). Aruna Perera (YT37) also points out that the artificial intelligence (AI) systems lack emotional intelligence and ability to learn from experience.

There was a shared belief that LLMs can make high-quality information available to a broader audience, breaking down traditional barriers to accessing expert advice. They offer opportunity to educate RIs and translate complex financial jargon into simple terms (YT25). Richard Walker reinforces this notion discussing HuggingGPT.

Hugging GPT can be trained to understand complex Financial Concepts, jargon, and market dynamics. By doing so, it can assist analysts in various tasks such as generating concise summaries of financial reports and news articles identifying market trends and patterns evaluating the credit worthiness of companies. (YT25)

LLMs also enhance the accessibility of expert opinions by democratising access to complex financial information, as one of the interviewees explained:

Experts are not accessible to most retail investors. Accessibility will lead to better understanding. (INT 1)

By improving access to information and providing educational to RIs, LLMs can empower them to make more informed decisions (YT3, YT4, YT5, YT11, YT21, YT33, YT42, YT49).

On the other hand, there were concerns voiced that the increased accessibility will not translate into improved understanding and decision-making by RIs. One interviewee stated

A retail investor does not want to know why the stock is outperforming; he just needs a name. (INT 4)

And another said

Increased understanding will come together with a higher risk of reliance on the model instead of using [one's] own brain capacity. (INT 3)

One respondent highlighted the concern that LLMs *fail to account for psychological factors* (INT 6) and may *misinterpret or mishandle information* (INT 4), posing potential risks to investors who might over-rely on AI-driven advice. 'Over-reliance' on LLMs could lead to unintended consequences in the investment process (YT41).

Increased accessibility may also not translate into more informed decision-making due to potential lack of nuance, hallucinations, incompleteness, or misuse of LLM-generated advice (for example, malicious advice) (YT1, YT5, YT17, YT29, YT34, YT35, YT37, YT40, YT44). As ine of the interviewees explained LLMs use can be bring the threat of *influencing and manipulating people (investors); there might be ill intentions and sabotage by LLM designers* (INT3) which could lead to errors in investment decisions. RIs were hesitant to trust LLMs as much as they trust human experts. As one of the interviewees noted *Retail investors at mass ..., uneducated will be suspicious. New technology. Less trust to LLM. More to human.* (INT 2) The concerns over reliability and accuracy (YT23) may require fact-checking of complex responses (YT1).

Despite this, some experts believe that trust in LLMs will grow as their capabilities improve and RIs will rely less on traditional brokers and fund managers (YTQ8). At the same time expert roles may grow as their advice and opinions would be readily available to a broader audience. Trust in LLM-generated advice may increase as AI models become more capable, reliable, and seamlessly integrated into user interfaces and workflows (YT1, YT2, YT4).

While some experts believe that LLMs will eventually replace certain functions traditionally performed by human experts, there is a consensus that their roles will evolve rather than disappear entirely. There was no expectation of pronounced changes in the roles of brokers, investment bankers and fund managers. Josh Brown pointed out that

Nobody with a straight face thinks a rich person wants to talk to a robot about his problems. (YT16)

However, this might change with the development of the technology, as one of the interviewees said:

The role of LLM will be ever-growing, with the potential to reach 65-90% of investment decision-making. Human experts will propel in the data acquiring, for example by interviewing the executives, visiting on-site operations etc – those operations which cannot be performed by LLMs. (YT6)

The most frequent opinion was that LLMs would enhance current roles, extending professionals abilities to deal with large volumes of data and provide more personalised service to clients (YT5, YT6, YT7, YT9, YT14, YT15, YT16). Robyn Grew underlined that financial professionals would need to use AI in their jobs to maintain competitive edge (YT6). One of the interviewed experts stated:

Their [financial intermediaries] role will be limited to 1) client interface and 2) servicing infrastructure such as back-office. In 2 years. Global back office of the capital market. (INT5)

Subjective Norms

SN refer to the opinions and behaviours of financial experts, influencers, and other authoritative figures in the investment community impacting retail investors` opinions. These opinions can influence RIs ' decisions, as one respondent explained

On the market I am familiar with the investors are mostly uneducated, so investors are heavily relying on speaking heads. (INT4)

This may be especially important when innovative technologies are considered. One of the interviewees underlined the importance of trust in human experts and the role of endorsements by well-known figures *"If it someone I trust, I will definitely consider LLM"* (INT3). Another interviewee expressed similar oponion *"Collectively discussed and approved recommendations have higher quality and lesser risks"* (INT2). One of the respondents highlighted the generational aspect, pointing out that older generations with lower trust in LLMs, might be swayed by endorsements from trusted figures (INT4). The endorsement of LLMs by respected financial experts could lead to greater acceptance and trust towards the new technology among RIs. A similar sentiment was expressed by another expert, who stated that:

Endorsement by experts will increase the trust of retail investors. This will help address older generations who have an inherently larger mistrust of LLM. (INT4)

Younger generations, who are more accustomed to digital tools, may increasingly view LLMs as credible sources of advice (INT4). This generational shift could eventually lead to broader acceptance and a redefinition of what is defined as expert advice in the investment community. As another expert suggested *Existing generations are mentally ready for adoption.* (INT2). However, some investors stated that successful cases and objective studies are more important for them than experts' endorsement (INT 2, INT 4).

The group norms are also expressed in the 'herding behaviour' among RIs, which may be amplified by LLMs endorsement and increased use by industry experts (INT 3). This, in turn, could lead to more uniform decision-making among investors, potentially reducing market inefficiencies but also increasing the risk of collective errors if the LLM-generated advice is flawed. It was also mentioned that widespread use of LLMs by RIs could increase market volatility (YT47). One of the interviewees explained

There could be a very wide range of possible impacts, from manipulation to higher volatility to super arbitrage and levelling out. New models and patterns of behaviour will emerge. (INT 5)

In several YT videos regulatory issues related to data ownership and privacy were mentioned (YT1, YT5, YT7, YT12, YT17, YT37, YT41). Regulatory frameworks should ensure the accuracy and reliability of AI outputs, and ethical considerations should ensure that human oversight remains a critical component of the investment process. Concerns were also raised regarding AI use leading to increased inequalities

Are they again just going to unlevel the playing field even more and make it available to those clients who have half a million or a million dollars to invest with them or are they going to do what we're trying to do in that is make the playing field level. (YT45)
and further

I don't think that's an improvement in our society when we just throw upon those who are very rich in more advantages over those who aren't. (YT45)

Perceived Behavioural Control

PBC refers to how RIs perceive their ability to use LLM-generated advice. The perception of how easy or difficult it is for RIs to access and use LLM-generated investment advice plays a critical role in determining how LLMs are used in retail investment.

One expert stated that *"It's not difficult and it should not be difficult to access"* LLM-generated advice and that *'it's just a matter of education and learning process'* (INT3). Others identified several challenges to implementing LLMs into retail investment such as *"lack of product awareness"* (INT6) and *'geographical limitations'* (INT4), which can hinder access to LLMs. The technical complexities involved in integrating LLMs into existing investment processes and required technical knowledge were also mentioned as significant challenges (YT1, YT2, YT5, YT6, YT11, YT13, YT14, YT17, YT21, YT51). As explained in one of the YT videos

There is this notion of prompt engineering, which is extremely important for you to understand so that you can leverage the most out of these models when you're wearing it and that's an engineering problem. (YT23)

The quality of the user interface (UI) and user experience (UX) also play role, as explained *"It all depends on the quality of UI/UX"* (INT7). Other barriers included the need for substantial financial resources, the complexity of data processing, and continuous learning and adaptation requirement (YT21).

LLMs have the potential to significantly enhance investors' confidence by providing tools that simplify complex financial information. As one of the experts explained "*More confidence, more feeling of control—LLM will empower retail investors"* (INT3). This was contrasted by concerns from another respondent who felt that the complexity of LLMs might reduce the perceived control; he said *"LLM is controlling me, due to my little understanding of the processes and factors"* (INT4). On the other hand, LLMs might create a false sense of control. As one expert suggested, "*you are seeing certain false facts come out as answers from Chat GPT which with a very confident sounding*" (YT23).

Scepticism and the difficulty in trusting AI-generated advice affect how comfortable RIs feel about relying on LLMs. One of the experts said that *"LLMs are intentionally limited in their capabilities to provide certain types of advice."* (INT4). Such believes could prevent widespread adoption of LLMs for investment purposes (YT1, YT9, YT17, YT23, YT29, YT34, YT42). This sentiment was echoed in several videos, which inferred that RIs do not trust LLM-generated advice as much as they trust advice from human experts (YT1, YT3, YT4).

The analysis of the YT video interviews brought up also challenges related to integrating LLMs with other technologies and integrating the tools into the existing investment decisions process. This requires the involvement of different market players and is both complex and resource-intensive (YT6). It also often remains outside of the users' control.

Usefulness of the AESTIMA tool for data extraction and analysis

The manual checks of data extraction and analysis described above showed 76% of correctly extracted data by subject and meaning (score 2) and 95% correctly extracted data by subject alone (score 1). This number was deemed sufficient as the results indicated consistent patterns within the tested scenarios, and additional tests did not yield further insights.

Discussion

RIs generally had a positive attitude toward LLMs appreciating their capabilities and recognising their limitations. Further development of the technology and a growing number of cases describing the successful use of LLMs in the retail investment field, the willingess to introduce them into professional practice should increase. However, failed LLMs applications, technical challenges, as well as regulatory and ethical issues related to the use of LLMs could hinder their implementation. The comparison between 2023 and 2024 reveals a gradual shift in sentiment towards greater acceptance and integration of LLMs in financial decision-making processes. While scepticism regarding trust and reliability remains, recognition of the potential benefits of LLMs also increases, particularly in enhancing investor confidence and complementing human expertise. Simplification of the tools and ongoing education are essential in overcoming barriers to the widespread use of LLMs in the investment industry.To fully benefit from LLMs benefits requires overcoming limitations, such as the need for more nuanced judgment and the inherent risks associated with reliance on AI-driven advice. Predicted democratisation of the industry could help RIs make more informed and more rational decisions. On the other hand, overreliance on the tools and underestimation or lack of understanding of their limitation could lead to decisions destabilising markets and resulting in loss for RIs.

Inaccuracies, misconceptions, and hallucinations generated by LLMs could lead to financial losses or harm to individuals' savings (Lo & Ross, 2024). Concerns regarding data privacy, bias, fairness, and transparency, risk of unethicall use of LLMs underscores the need for robust safeguards and ethical guidelines (Petroșanu et al., 2023; Sun et al., 2024). The combination of ‘word of mouth’ recommendations and high-level regulation would shape the norms around using LLMs within the industry .The EU Artificial Intelligence Act (*EU Artificial Intelligence Act*, 2024) and the General Data Protection Regulation (GDPR) (*General Data Protection Regulation (GDPR) – Legal Text*, 2018) emphasise the need for compliance with strict data privacy and ethical standards (Chen et al., 2024; Leslie & Perini, 2024). Wealthier regions and individuals who can afford to comply with stringent regulations or who are in less regulated regions could benefit more from the advancements in AI.

There seems to be an agreement that those who would embrace new technology will gain an advantage in their areas of expertise. Widespread use of the LLMs within the sector could eventually lead to a partial replacement of human experts in certain areas, and a change of roles of retail investment experts may be expected. There is, however, no imminent risk of their complete replacement with technology, as at this stage of its development, it does not offer reliability sufficient to gain complete trust from the investors. Exploring how old roles evolve and what new roles will be created should be focus on future research to inform the education of those working in the financial sector as well as guide the systematic approach to offset potential job loss or necessary re-skilling (Leslie & Perini, 2024).

The TPB has proven to be a useful analysis framework, guiding its focus and allowing in-depth exploration of the topic. Core concepts of the theory were easily identifiable in the sources and painted a picture of the current and predicted state of intentions to use LLMs in RIs decisions making process. The use of the AESTIMA tool allowed the first stage of analysis to be completed fast and with high level of accuracy. This facilitated an engagement with large volume of data, which would be more resource intensive in a traditional approach. This approach could be particularly usefy for smaller teams and/or teams with limited resources. The insights generated when using the tool were comprehensive and were well placed within the wider context. It is important to note that the analysis was completed on a curated set of collected sources, which removes the risk of hallucinations of non-existing sources frequently cited in the literature (Hadi et al., 2023; Huang et al., 2023). Additionally, the human input was still at the core of analysis, from checking the accuracy to the interpretation of the AESTIMA tool output. However, at this stage of the technology development it should not replace the researchers, as their knowledge of the topic, engagement and familiarisation with the sources are still essential. Use of any of the LLMs based tools in research process should be undertaken with knowledge of advantages and limitations of the technology and follow the ethical guidance (Porsdam Mann et al., 2024).

The findings are potentially applicable to various sectors of the financial industry. In asset management, LLMs can optimise portfolio strategies by analysing market trends and assessing risk more efficiently than human experts. Similarly, in regulatory compliance, LLMs can streamline the interpretation of complex regulatory documents, enhance audit processes, and improve anomaly detection for fraud prevention. In wealth management, LLMs can deliver personalised investment advice by aggregating data across client portfolios, aligning well with the industry’s focus on tailored financial planning. These applications are particularly relevant due to the financial industry’s reliance on regulation, data accuracy, and customer protection. The capability to democratise access to complex tools can also address systemic inefficiencies of other sectors, making expert-level insights available to underserved populations. However, as noted, ethical concerns like over-reliance, manipulation, and inherent model biases require robust oversight, which is critical to all sectors where trust and compliance are non-negotiable.

The limitation of this study can be used to formulate direction of future reseaerch. The interviewees were recruited through the personal network of the PI, who is a retail investor himself, this could introduce a selection bias. However, their opinions were matched with those from the identified YT videos, which may indicate that they reflect commonly shared sentiments well. A selection of YT video transcripts was analysed due to existing large volume of sources. This may lead to some less frequent views being missed in the analysis. Only videos in English were included in the analysis, which can introduce further bias, especially considering a notable lack of voices from certain parts of the world, such as Africa or South America. Inclusion of varied voices from YT video recording, addressed to some extent lack of wider demographic respresentation in the interviews. Further research should also focus on exploring the impact on LLMs on RIs behaviour, encompassing worldwide perspectives, reflecting the interconnected nature of financial markets and consider issues of diversity in the financial industry.

Conclusions

LLMs have the potential to change the retail investment industry, influencing professional roles within, impacting who has access to it, and changing market dynamics. Despite great promises, caution is needed when implementing LLMs due to their limitations, as is the case for any innovation at its early stages of development. At the current state, LLMs cannot replace human experts, but there is a potential for that. Education and endorsement by well-established experts in the field remain essential for the successful large-scale adoption of LLMs. Human experts continue to play a role in interpreting and validating AI-generated insights, ensuring that the advice provided to RIs is sound and reliable. Integration of LLMs into financial advisory services should be accompanied by robust regulatory frameworks and ethical guidelines to ensure that they serve to enhance, rather than detract from, the quality of investment advice available to RIs. LLMs can potentially enhance accessibility to industry expert insights for retail investors, democratising investment advice and improving informed decision-making. They can also help to reduce emotional biases by offering data-driven insights, while simplifying complex financial information for increased understanding. However, the results also caution against over-reliance on LLMs due to their current lack of nuance and potential for errors.

Acknowledgement

Author Contributions: Conceptualization, D.G., M.G., A.B., S.K., V.A, I.L.; Methodology, D.G., M.G., A.B., S.K., V.A, I.L.; Data Collection, D.G., M.G., A.B., S.K., V.A, I.L.; Formal Analysis, D.G., M.G., A.B., S.K., V.A, I.L.; Writing—Original Draft Preparation, D.G., M.G., A.B., S.K., V.A, I.L.; Writing—Review and Editing, D.G., M.G., A.B., S.K., V.A, I.L. All authors have read and agreed to the published the final version of the manuscript.

Institutional Review Board Statement: Ethical review and approval were waived for this study, due to that the research does not deal with vulnerable groups or sensitive issues.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy.

Conflicts of Interest: The authors declare no conflict of interest.

References

Adam, A. A., & Shauki, E. R. (2014). Socially responsible investment in Malaysia: Behavioral framework in evaluating investors' decision making process. Journal of Cleaner Production, 80, 224–240. https://doi.org/10.1016/j.jclepro.2014.05.075

Ajzen, I. (2011). The theory of planned behaviour: Reactions and reflections. Psychology & Health, 26, 1113–1127. https://doi.org/10.1080/08870446.2011.613995

Bikhchandani, S., & Sharma, S. (2000). Herd behavior in financial markets. IMF Staff Papers, 47(3), 279–310. https://doi.org/https;//doi.org/10.2307/3867650

Braun, V., & Clarke, V. (2022). Conceptual and design thinking for thematic analysis. Qualitative Psychology, 9(1), 3.

Campbell, S. D., & Sharpe, S. A. (2009). Anchoring bias in consensus forecasts and its effect on market prices. Journal of Financial and Quantitative Analysis, 44(2), 369–390. https://doi.org/10.1017/S0022109009090127

Cao, H. (2024). Recent advances in text embedding: A Comprehensive Review of Top-Performing Methods on the MTEB Benchmark. arXiv Preprint arXiv:2406.01607.

Cao, Z., & Feinstein, Z. (2024). Large Language Model in Financial Regulatory Interpretation. arXiv Preprint arXiv:2405.06808.

Chen, Z. Z., Ma, J., Zhang, X., Hao, N., Yan, A., Nourbakhsh, A., Yang, X., McAuley, J., Petzold, L., & Wang, W. Y. (2024). A Survey on Large Language Models for Critical Societal Domains: Finance, Healthcare, and Law. arXiv Preprint arXiv:2405.01769.

Clarke, V., & Braun, V. (2013). Successful qualitative research: A practical guide for beginners. Sage publications ltd.

Daudert, T. (2021). Exploiting textual and relationship information for fine-grained financial sentiment analysis. Knowledge-Based Systems, 230, 107389. https://doi.org/10.1016/j.knosys.2021.107389

Dewi, M. L., & Ronny, R. (2023). The Effect of Theory of Planned Behavior and Customer Relationship Marketing on Mutual Fund Investment Intentions. Jemasi: Jurnal Ekonomi Manajemen Dan Akuntansi, 19(1), 87–102. https://doi.org/10.35449/jemasi.v19i1.653

Dierkes, M., Klos, A., & Langer, T. (2009). Do my friends influence my investment behavior? Evidence from a representative sample of the German population. European Financial Management Association 2010 Annual Meetings, Aarhus, Denmark. https://www.efmaefm.org/0efmameetings/efma%20annual%20meetings/2010-Aarhus%20old/EFMA2010_0335_fullpaper.pdf

EU Artificial Intelligence Act. (2024). https://artificialintelligenceact.eu/

Gamel, J., Bauer, A., Decker, T., & Menrad, K. (2022). Financing wind energy projects: An extended theory of planned behavior approach to explain private households' wind energy investment intentions in Germany. Renewable Energy, 182, 592–601. https://doi.org/10.1016/j.renene.2021.09.108

Gao, Y., Xiong, Y., Gao, X., Jia, K., Pan, J., Bi, Y., Dai, Y., Sun, J., Guo, Q., Wang, M., & Wang, H. (2024). Retrieval-Augmented Generation for Large Language Models: A Survey (No. arXiv:2312.10997). arXiv. http://arxiv.org/abs/2312.10997

General Data Protection Regulation (GDPR) – Legal Text. (2018). https://gdpr-info.eu/

Gimmelberg, D., Glowacka, M., Belinskiy, A., Korotkii, S., Artamov, V., & Ludviga, I. (n.d.). Use of Large Language Models in the asset management industry – evidence review. Forthcoming.

Gopi, M., & Ramayah, T. (2007). Applicability of theory of planned behavior in predicting intention to trade online: Some evidence from a developing country. International Journal of Emerging Markets, 2(4), 348–360. https://doi.org/10.1108/17468800710824509

Gumasing, M. J. J., & Niro, R. H. A. (2023). Antecedents of Real Estate Investment Intention among Filipino Millennials and Gen Z: An Extended Theory of Planned Behavior. Sustainability, 15(18), 13714. https://doi.org/10.3390/su151813714

Guo, Z., Jiang, G., Zhang, Z., Li, P., Wang, Z., & Wang, Y. (2023). Shai: A large language model for asset management. arXiv Preprint arXiv:2312.14203.

Hadi, M. U., Al-Tashi, Q., Qureshi, R., Shah, A., Muneer, A., Irfan, M., Zafar, A., Shaikh, M., Akhtar, N., Wu, J., & Mirjalili, S. (2023). Large Language Models: A Comprehensive Survey of its Applications, Challenges, Limitations, and Future Prospects. https://doi.org/10.36227/techrxiv.23589741.v1

Huang, L., Yu, W., Ma, W., Zhong, W., Feng, Z., Wang, H., Chen, Q., Peng, W., Feng, X., & Qin, B. (2023). A survey on hallucination in large language models: Principles, taxonomy, challenges, and open questions. arXiv Preprint arXiv:2311.05232.

Islamoğlu, M., Apan, M., & Ayvalı, A. (2015). Determination of factors affecting individual investor behaviours: A study on bankers. International Journal of Economics and Financial Issues, 5(2), 531–543.

Keraghel, I., Morbieu, S., & Nadif, M. (2024). Beyond words: A comparative analysis of LLM embeddings for effective clustering. Miliou, I., Piatkowski, N., Papapetrou, P. (Eds) Advances in Intelligent Data Analysis XXII. IDA 2024. Lecture Notes in Computer Science, 14641, 205–216. https://doi.org/10.1007/978-3-031-58547-0_17

Kisaka, S. E. (2014). The Impact of Attitudes towards Saving, Borrowing and Investment on the Capital Accumulation Process in Kenya: An Application of the Theory of Planned Behavior. Research Journal of Finance and Accounting, 5(9), 140–151.

Leslie, D., & Perini, A. M. (2024). Future Shock: Generative AI and the international AI policy and governance crisis. Harvard Data Scence Review, Special Issue 5. https://doi.org/10.1162/99608f92.88b4cc98

Li, Y., Wang, S., Ding, H., & Chen, H. (2023). Large language models in finance: A survey. 374–382. https://doi.org/10.1145/3604237.3626869

Lo, A. W., & Ross, J. (2024). Can ChatGPT plan your retirement?: Generative AI and financial advice. Generative AI and Financial Advice (February 11, 2024).

Ma, F., Lyu, Z., & Li, H. (2024). Can ChatGPT predict Chinese equity premiums? Finance Research Letters, 105631.

Mahalakshmi, T., & Anuradha, N. (2018). Factors affecting Investment decision making & investment performance among individual investors in India. International Journal of Pure and Applied Mathematics, 118(18), 1667–1675.

Malzara, V. R. B., Widyastuti, U., & Buchdadi, A. D. (2023). Analysis of Gen Z's Green Investment Intention: The Application of Theory of Planned Behavior. Jurnal Dinamika Manajemen Dan Bisnis, 6(2), 63–84. https://doi.org/10.21009/JDMB.06.2.5

Neuendorf, K. A. (2018). Content analysis and thematic analysis. In Advanced research methods for applied psychology (pp. 211–223). Routledge. https://www.taylorfrancis.com/chapters/edit/10.4324/9781315517971-21/content-analysis-thematic-analysis-kimberly-neuendorf

Norisnita, M., & Indriati, F. (2022). Application of Theory of Planned Behavior (TPB) in Cryptocurrency Investment Prediction: A Literature Review. Economics and Business Quarterly Reviews, 5(2), 181–188. https://doi.org/10.31014/aior.1992.05.02.424

Nugraha, B. A., & Rahadi, R. A. (2021). Analysis of Young Generations toward Stock Investment Intention: A Preliminary Study in an Emerging Market. Journal of Accounting and Investment, 22(1), 80–103. https://doi.org/10.18196/jai.v22i1.9606

Nugraha, V. A., & Prasetyaningtyas, S. W. (2023). Analysis of Factors Influencing Investment Intention in Cryptocurrency: A Theory of Planned Behavior (TPB) Approach. Jurnal Ekonomi, 12(02), 541–551. https://ejournal.seaninstitute.or.id/index.php/Ekonomi/article/view/1653

Pahlevi, R. W., & Oktaviani, I. I. (2018). Determinants of Individual Investor Behaviour in Stock Investment Decisions. Accounting and Financial Review, 1(2), 53–61. https://doi.org/10.26905/afr.v1i2.2427

Patil, S., & Bagodi, V. (2021). A study of factors affecting investment decisions in India: The KANO way. Asia Pacific Management Review, 26(4), 197–214. https://doi.org/10.1016/j.apmrv.2021.02.004

Petroșanu, D.-M., Pîrjan, A., & Tăbușcă, A. (2023). Tracing the Influence of Large Language Models across the Most Impactful Scientific Works. Electronics, 12(24), 4957. https://doi.org/10.3390/electronics12244957

Petukhova, A., Matos-Carvalho, J. P., & Fachada, N. (2024). Text clustering with LLM embeddings. arXiv Preprint arXiv:2403.15112.

Pham, Q. T., Phan, H. H., Cristofaro, M., Misra, S., & Giardino, P. L. (2021). Examining the Intention to Invest in Cryptocurrencies: An Extended Application of the Theory of Planned Behavior on Italian Independent Investors. International Journal of Applied Behavioral Economics (IJABE), 10(3), 59–79. https://doi.org/10.4018/IJABE.2021070104

Piyumini, T. B., & Wijethunga, A. W. G. C. N. (2020). Effect of investors' psychology on capital market investment: An application of the theory of planned behavior. Journal of Management Matters, 7(1), 1–10. http://repository.rjt.ac.lk/handle/123456789/5037

Porsdam Mann, S., Vazirani, A. A., Aboy, M., Earp, B. D., Minssen, T., Cohen, I. G., & Savulescu, J. (2024). Guidelines for ethical use and acknowledgement of large language models in academic writing. Nature Machine Intelligence, 6, 1272–1274.

Raiaan, M. A. K, Mukta, M. S. H., K. Fatema, Fahad, N. M., S. Sakib, M. M. J. Mim, J. Ahmad, M. E. Ali, & S. Azam. (2024). A Review on Large Language Models: Architectures, Applications, Taxonomies, Open Issues and Challenges. IEEE Access, 12, 26839–26874. https://doi.org/10.1109/ACCESS.2024.3365742

Rooh, S., El-Gohary, H., Khan, I., Alam, S., & Shah, S. M. A. (2023). An Attempt to Understand Stock Market Investors' Behaviour: The Case of Environmental, Social, and Governance (ESG) Forces in the Pakistani Stock Market. Journal of Risk and Financial Management, 16(12). Scopus. https://doi.org/10.3390/jrfm16120500

Sari, R., Kusnanto, K., & Aswindo, M. (2022). Determinants of Stock Investment Decision Making: A Study on Investors in Indonesia. Golden Ratio of Finance Management, 2(2), 120–131. https://doi.org/10.52970/grfm.v2i2.174

Schifter, D. E., & Ajzen, I. (1985). Intention, perceived control, and weight loss: An application of the theory of planned behavior. Journal of Personality and Social Psychology, 49(3), 843–851. https://doi.org/10.1037/0022-3514.49.3.843

Setyorini, N., & Indriasari, I. (2020). Does millennials have an investment interest? Theory of planned behaviour perspective. Diponegoro International Journal of Business, 3(1), 28–35. https://doi.org/1597077849

Sharma, M., & Gupta, S. (2011). Role of Subjective Norm in Investment Decision Making of Casual Investors. Indian Journal of Finance, 5(11), 39–46. http://gnanaganga.inflibnet.ac.in:8080/jspui/handle/123456789/8005

Shefrin, H., & Statman, M. (2000). Behavioral portfolio theory. Journal of Financial and Quantitative Analysis, 35(2), 127–151. https://doi.org/doi:10.2307/2676187

Subramanian, Y. R. (2021). Social-Media Influence on the Investment Decisions Among the Young Adults in India. Advancement in Management and Technology (AMT), 2(1), 17–26. https://doi.org/10.46977/apjmt.2021v02i01.003

Sun, L., Huang, Y., Wang, H., Wu, S., Zhang, Q., Gao, C., Huang, Y., Lyu, W., Zhang, Y., Li, X., Liu, Z., Liu, Y., Wang, Y., Zhang, Z., Kailkhura, B., Xiong, C., Xiao, C., Li, C., Xing, E., … Zhao, Y. (2024). TrustLLM: Trustworthiness in Large Language Models (No. arXiv:2401.05561). arXiv. http://arxiv.org/abs/2401.05561

Vaismoradi, M., Jones, J., Turunen, H., & Snelgrove, S. (2016). Theme development in qualitative content analysis and thematic analysis. Journal of Nursing Education and Practice, 6, 100–110. https://doi.org/10.5430/jnep.v6n5p100

Widyastuti, U., Febrian, E., Sutisna, S., & Fitrijanti, T. (2023). Could the Theory of Planned Behaviour Explain Market Discipline in Sharia Mutual Funds? Australasian Accounting, Business and Finance Journal, 17(4), 3–20. Scopus. https://doi.org/10.14453/aabfj.v17i4.02

Yeo, K. H. K., Lim, W. M., & Yii, K.-J. (2023). Financial planning behaviour: A systematic literature review and new theory development. Journal of Financial Services Marketing. Scopus. https://doi.org/10.1057/s41264-023-00249-1

Yu, W. (2022). Retrieval-augmented generation across heterogeneous knowledge. Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies: Student Research Workshop, 52–58. https://doi.org/10.18653/v1/2022.naacl-srw.7

Yulandreano, E., & Rita, M. R. (2023). Investment Decisions on the Crowdfunding Platform Based on the Theory of Planned Behavior. Jurnal Manajemen Bisnis, 14(1), 36–52. https://doi.org/10.18196/mb.v14i1.16494

Zhao, P., Zhang, H., Yu, Q., Wang, Z., Geng, Y., Fu, F., Yang, L., Zhang, W., & Cui, B. (2024). Retrieval-augmented generation for ai-generated content: A survey. arXiv Preprint arXiv:2402.19473.

Zhao, X., Lu, J., Deng, C., Zheng, C., Wang, J., Chowdhury, T., Yun, L., Cui, H., Xuchao, Z., & Zhao, T. (2023). Domain specialization as the key to make large language models disruptive: A comprehensive survey. arXiv Preprint arXiv:2305.18703. https://www.researchgate.net/profile/Chen-Ling-30/publication/371163915_Domain_Specialization_as_the_Key_to_Make_Large_Language_Models_Disruptive_A_Comprehensive_Survey/links/64c805568b5de83215d9fc39/Domain-Specialization-as-the-Key-to-Make-Large-Language-Models-Disruptive-A-Comprehensive-Survey.pdf.