

Review

Not peer-reviewed version

---

# A Comprehensive Review of Generative AI in Finance

---

David Kuo Chuen Lee, [Chong Guan](#), [Yinghui Yu](#), [Qinxu Ding](#)\*

Posted Date: 26 July 2024

doi: 10.20944/preprints202407.2109.v1

Keywords: Generative AI; Large Language Models; Finance; Topic Modeling



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

*Review*

# A Comprehensive Review of Generative AI in Finance

David Kuo Chuen Lee, Chong Guan, Yinghui Yu and Qinxu Ding \*

Singapore University of Social Sciences

\* Correspondence: qinxuding@suss.edu.sg

**Abstract:** The integration of generative AI (GAI) into the financial sector has brought about significant advancements, offering new solutions for various financial tasks. This review paper provides a comprehensive examination of recent trends and developments at the intersection of GAI and finance. By utilizing an advanced topic modeling method, BERTopic, we systematically categorize and analyze existing research to uncover predominant themes and emerging areas of interest. Our findings reveal the transformative impact of finance-specific large language models (LLMs), the innovative use of generative adversarial networks (GANs) in synthetic financial data generation, and the pressing necessity of a new regulatory framework to govern the use of GAI in the finance sector. This paper aims to provide researchers and practitioners with a structured overview of the current landscape of GAI in finance, offering insights into both the opportunities and challenges presented by these advanced technologies.

**Keywords:** generative AI; large language models; finance; topic modeling

## 1. Introduction

The intersection of generative AI (GAI) and finance has emerged as a rapidly developing area of research and application, revolutionizing various facets of the financial industry. GAI encompasses a broad range of models, such as Variational Autoencoders (VAEs), Generative Adversarial Networks (GANs), Large Language Models (LLMs), and Diffusion models. It has demonstrated significant potential in enhancing financial analytics, improving decision-making processes, and generating synthetic financial data for various applications. This review paper aims to provide a comprehensive overview of recent trends and advancements in the GAI's application within the financial sector.

The field of GAI has experienced rapid advancement in recent years. One of the famous GAI methods is the VAE algorithm, developed by Kingma and Welling in 2013 [1]. VAEs applied probabilistic frameworks to generate new data points based on latent representations learned during training. Therefore, it enhances the ability to model complex distributions in data. Subsequently, in 2014, Goodfellow et al. [2] introduced the GAN model, which represented a paradigm shift in generative modeling. It is an adversarial process involving two neural networks: a generative model to capture the data distribution and a discriminative model to distinguish between real data and generated samples. This adversarial framework enabled GANs to produce remarkably realistic outputs across various domains, such as images and texts.

In 2017, Transformers was introduced by Vaswani et al. [3]. It leverages self-attention mechanisms to capture long-range dependencies in sequential data, making them highly effective for various natural language tasks. This innovation inspired the development of LLMs, particularly transformer-based architectures like OpenAI's GPT series [4–7] and Google's BERT [8], which marked a significant advancement in natural language processing. The deployment of LLMs has expanded generative AI into sophisticated applications requiring the comprehension and generation of human-like text. Additionally, Diffusion models have emerged as a novel approach in GAI, focusing on modeling temporal dependencies and irregular patterns in sequential data [9]. These models have shown promise in applications where traditional generative models fall short, particularly in capturing the nuances of dynamic and time-series data.

From the financial researchers' perspective, the impact of LLMs applications is most significant and far-reaching. In the past two decades, financial text mining has been a popular research area,

especially with advancements of computational methods that have made processing large-scale data possible. Beyond conventional financial data sources such as public companies' annual reports and earning announcements, financial researchers have turned to financial news press, regulatory filings, and social media to uncover hidden information and sentimental cues. These insights can be used to predict investment behaviors and trends in stock returns. For instance, Bollen et al. [10] investigated whether measurements of collective mood states derived from large-scale Twitter feeds are correlated to the value of the Dow Jones Industrial Average (DJIA) over time. Wisniewski and Yekini [11] analyzed the qualitative part of annual reports of UK-listed companies and used the frequency of words associated with different language indicators to forecast future stock returns. McGurk et al. [12] examined the relationship between investor sentiment and stock returns by employing textual analysis on Twitter posts and found that their investor sentiment measure has a positive and significant effect on abnormal stock returns.

Most of these earlier studies are still exploratory, often reducing text information to "a bag of words" or representing it through dictionary-based sentiment scores. This is primarily because financial texts frequently lack a regular structure, and verbal/textual communication can be subtle and complex. Additionally, financial jargon can vary in meaning depending on the context. To extract more meaningful insights from financial texts, more advanced models are required. LLMs trained on extensive datasets from diverse sources and themes can provide more sophisticated text representations that capture the nuances of financial language. In the literature we surveyed, Gupta [13] simplified the process of assessing annual reports of all the firms by leveraging the capabilities of LLMs, where the insights generated by the LLM are compiled in a Quant styled dataset and augmented by historical stock price data. Fatemi et al [14] showcased the remarkable capabilities of LLMs, even smaller models, in both fine-tuning and in-context learning for financial sentiment analysis. Li et al. [15] reported evidence that general-purpose LLMs especially GPT-4 could outperform domain-specific models in terms of sentiment analysis. Pavlyshenko [16] demonstrated that Llama 2 can be fine-tuned and multitask – when analyzes financial texts it can return both a structured response and sentiment data in specified JSON format, which can further be loaded directly into predictive models as features. On the other hand, Xing [17] reported that a design framework with heterogeneous LLM agents can be effective in financial sentiment analysis without fine-tuning.

In general, the financial sector, characterized by its vast and complex data, stands to benefit immensely from these advancements in GAI. Traditional data analysis methods often fall short in handling the scale, variability, and intricate patterns inherent in financial data. GAI offers a promising solution by not only managing large datasets effectively but also generating synthetic data close to real-world financial data. This capability is particularly important for applications such as risk management, fraud detection, algorithmic trading, and financial forecasting. Recent years have seen the emergence of specialized GAI models tailored for financial applications. Finance-specific LLMs, such as FinGPT [18,19] and FinPT [20], have been developed to address domain-specific challenges and have shown superior performance compared to general-purpose models in various financial tasks. Despite the promising advancements, the integration of GAI in finance is not without challenges. Issues such as data privacy, model interpretability, regulatory compliance, and the potential for generating biased or misleading data necessitate careful consideration. The ethical and social implications of deploying generative AI in financial decision-making further underscore the need for robust frameworks and guidelines to ensure responsible use [21].

Recently, Ding et al. [22] conducted an extensive examination of LLMs. Li et al. [23] and Lee et al. [24] provided surveys specifically focused on the impact of LLMs in finance. Additionally, Barde and Kulkarni [25], Krause [26], and Mbanyele [27] concentrated on general-purpose LLMs, such as ChatGPT, Bard, and Bing AI. However, our paper broadens the scope to explore the intersection of generative AI (GAI) and finance, extending beyond just LLMs. Furthermore, we employ an advanced topic modeling method, BERTopic [28], to comprehensively cluster and analyze the existing body of

work. By leveraging the BERTopic model, we present a novel framework for classifying the existing research on GAI and finance.

The structure of our paper is as follows: In Section 2, we introduce the dataset and the topic modeling method used in this study. Section 3 presents the results of our analysis. In Section 4, we provide an in-depth discussion based on the new framework obtained through the topic modeling method. Finally, in Section 5, we conclude our work and discuss future directions and potential areas for further research.

## 2. Materials and Methods

This study leverages the Google Scholar database for its extensive coverage, interdisciplinary reach, and up-to-date research indexing. Given that research on generative artificial intelligence (GAI), particularly large language models (LLMs), is relatively new and rapidly evolving, Google Scholar's comprehensive indexing of reputable authors and institutions is particularly valuable. Many significant papers are submitted to repositories such as arXiv and SSRN, both of which are well indexed by Google Scholar.

We conducted searches using two key combinations: (1) "generative AI and finance" and (2) "large language models and finance." From these searches, we retrieved a total of 90 papers published between 2018 and 2024. These papers were sourced from a diverse range of publishers, including ACM, ACL, arXiv, Curran Associates, Darcy & RoyPress, Elsevier, IEEE, MDPI, Routledge, SSRN, Taylor & Francis Online, and MIT Press.

To effectively cluster and analyze these papers, we applied a robust topic modeling technique known as BERTopic [28]. Topic models are powerful unsupervised tools for uncovering themes and underlying narratives in textual data. While traditional topic modeling methods such as Latent Dirichlet Allocation (LDA) [29] and Non-Negative Matrix Factorization (NMF) [30] represent documents as mixtures of latent topics using a bag-of-words approach, BERTopic enhances this process with advanced techniques.

To generate coherent topic representations, the BERTopic model was employed with clustering techniques and a class-based variation of Term Frequency-Inverse Document Frequency. Furthermore, BERTopic utilizes a pretrained LLM to create document embeddings, significantly improving the quality and coherence of the generated topics. BERTopic has shown effectiveness in various applications, including systematic reviews [13,22]. Our application of BERTopic in this study can be broken down into the following steps:

- **Data Preprocessing:** first, we convert all text to lowercase to ensure uniformity and reduce redundancy; second, we use `nltk.word_tokenize()` to split the text into individual tokens and `WordNetLemmatizer()` to reduce words to their base or root form; finally, we use `stopwords.words('english')` to eliminate the common stopwords as they usually do not contribute significantly to the meaning of text.
- **Fit the Model and Transform Documents:** we use *BERTopic* and *ClassIFidTransformer* to fit the model to our data and transform to discover topics.
- **Topics Exploration:** after fitting the model, we explore the topics generated by various tools of BERTopic.

## 3. Results

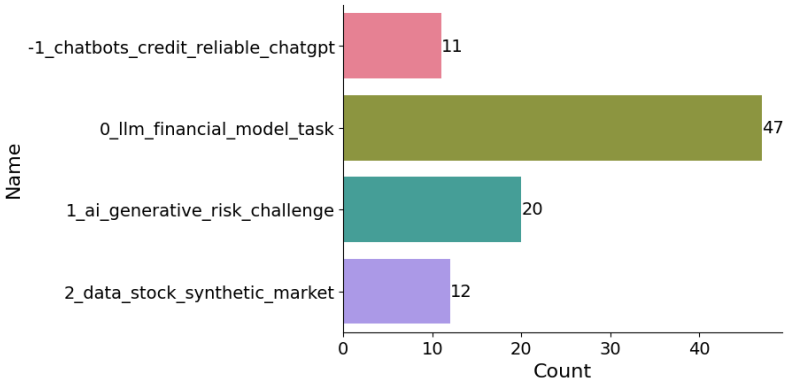
In this section, we discuss the results obtained from the BERTopic model. After fitting the model, we identified the most frequent topics in our dataset, as shown in Figure 1. The Figure 1 reveals four distinct clusters:

- "-1\_chatbots\_credit\_reliable\_chatgpt"
- "0\_llm\_financial\_model\_task";
- "1\_ai\_generative\_risk\_challenge";

- and "2\_data\_stock\_synthetic\_market".

Cluster -1 represents all outliers and should be disregarded. Consequently, our focus will be directed towards the examination of the remaining three clusters. Cluster 0 pertains to discussions surrounding the application of LLMs in financial tasks. This cluster highlights the innovative use of LLMs to address various financial modeling and task automation challenges. Cluster 1 delves into the challenges and risks associated with the implementation of GAI within the realm of finance. This cluster underscores the potential risks and regulatory considerations that accompany the deployment of GAI technologies in finance. Cluster 2 centers on the generation of synthetic financial data facilitated by GAI. This cluster emphasizes the role of GAI in creating synthetic datasets, which are essential for tasks such as market simulation and risk assessment.

Additionally, our analysis reveals a significant discrepancy in the distribution of research focus. Specifically, there are approximately 47 papers discussing LLMs in finance, a considerably higher count compared to those addressing the risks and data generation aspects of GAI. This observation underscores that LLMs currently represent the foremost research focus of GAI within the financial domain.



**Figure 1.** This figure describes the frequent topics obtained from the BERTopic model.

Observing Table 1, we identify the representative words for each topic generated by the BERTopic model. These words help elucidate the main themes and concepts of each topic, as they are extracted based on their relevance and frequency within the topic. Topic -1 is ignored as it is an outlier, so our analysis begins with Topic 0: "LLMs for Financial Tasks."

**Topic 0: "LLMs for Financial Tasks":** In addition to the words 'llm', 'financial', and 'task' that appear in the topic name, words such as 'model', 'language', and 'large' are also prominent, highlighting the characteristics of LLMs. Since LLMs are a type of language model, the data they process is human 'text', often requiring specific 'instructions' for interaction. The terms 'benchmark' and 'performance' relate to evaluating the effectiveness of LLMs in executing financial tasks.

**Topic 1: "The Risk and Challenge of Generative AI":** The words 'ai', 'artificial', 'intelligence', 'generative', 'risk', and 'challenge' appear in the topic name. Additionally, the term 'ethical' frequently emerges, underscoring significant concerns regarding the application of GAI in finance. These ethical issues are often discussed within the 'industry'. The word 'paper' is common in our dataset due to the review nature of this article but does not hold significant meaning in this context.

**Topic 2: "Synthetic Financial Data Generation":** The words 'synthetic' and 'data' appear in the topic name, while terms such as 'network', 'GANs', 'learning', and 'adversarial' indicate that most existing papers utilize Generative Adversarial Networks (GANs) to generate synthetic 'stock' 'market' data. This data type, consisting of 'price' 'series', is crucial for financial analyses and simulations.



**Table 1.** This table includes the representation words of each topic generated by BERTopic model.

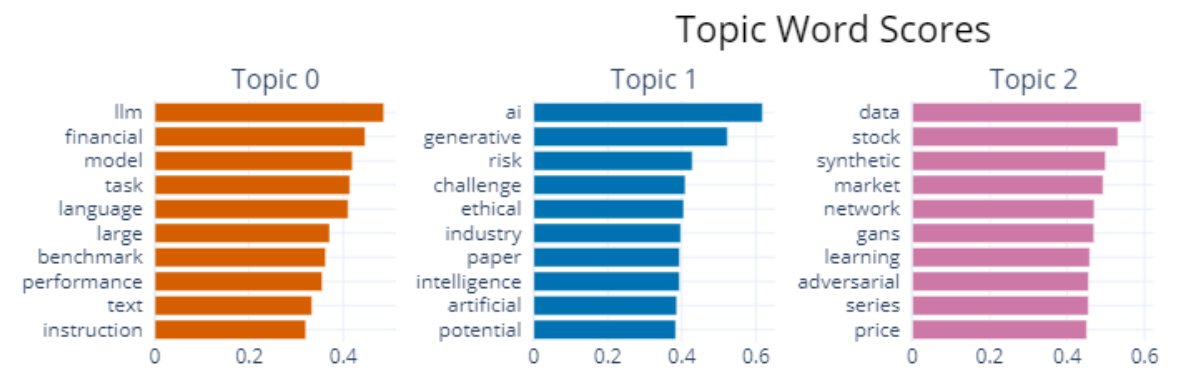
Count	Name	Representation
11	-1_chatbots_credit_reliable_chatgpt	['chatbots', 'credit', 'reliable', 'chatgpt', 'lgp', 'payment', 'user', 'transaction', 'individual', 'process']
47	0_llm_financial_model_task	['llm', 'financial', 'model', 'task', 'language', 'large', 'benchmark', 'performance', 'text', 'instruction']
20	1_ai_generative_risk_challenge	['ai', 'generative', 'risk', 'challenge', 'ethical', 'industry', 'paper', 'intelligence', 'artificial', 'potential']
12	2_data_stock_synthetic_market	['data', 'stock', 'synthetic', 'market', 'network', 'gans', 'learning', 'adversarial', 'series', 'price']

Figure 2 illustrates the word scores for each topic, indicating the importance or relevance of each word to its specific topic. These scores help us understand the strength of association between a word and a specific topic. As discussed by Grootendorst [28] in the BERTopic model, the c-TF-IDF method is used to calculate the word score for each cluster or topic. The definition of c-TF-IDF is as follows:

$$W_{x,c} = \|tf_{x,c}\| \times \log(1 + \frac{A}{f_x}),$$

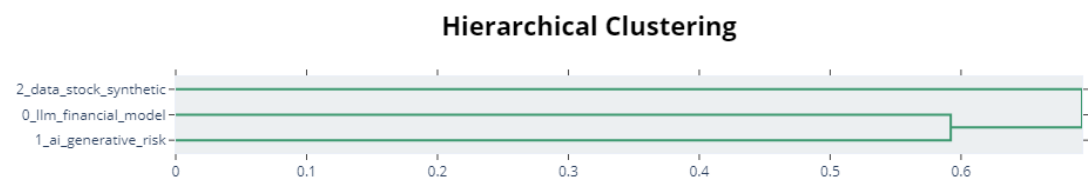
(1)

where x is the word, c is the cluster,  $tf_{x,c}$  is the frequency of word x in cluster c,  $f_x$  refers to the word x's frequency across all clusters and A represents the word's average number per cluster. Figure 2 tells us the most important 10 words for each cluster and their importance. For example, the word "llm" is most important for topic 0.



**Figure 2.** This figure shows the word scores for each topic.

Figure 3 presents the dendrogram of the hierarchical clustering of the three clusters. The x-axis represents the distance or dissimilarity between clusters, while the vertical lines indicate the points at which clusters are merged. For instance, clusters 0 and 1 merge at a dissimilarity level of approximately 0.6, indicating their relative similarity compared to cluster 2. This visualization helps us understand the relationships and similarities among the identified topics.



**Figure 3.** This figure describes the hierarchical clustering result obtained from the BERTopic model.

## 4. Discussion

In this section, we delve into three key topics derived from the BERTopic analysis. The first topic explores the application of Large Language Models (LLMs) in finance across various tasks. This includes discussions on the capability of general-purpose LLMs (e.g., GPTs, Gemini) to address financial problems, the effectiveness of finance-specific LLMs (e.g., FinGPT, FinPT) compared to general-purpose LLMs, and the identification of benchmarks and financial datasets that can be used to fairly evaluate the performance of LLMs in finance. The second topic addresses the potential risks and challenges associated with using generative AI (GAI) models for financial applications. This includes an examination of issues such as hallucinations, ethical and social impacts, and financial regulation. Finally, the third topic focuses on the use of GAI for synthetic financial data generation. We will discuss the challenges and areas of focus in this domain, as well as existing work utilizing Variational Autoencoders (VAEs), Generative Adversarial Networks (GANs), and Diffusion Models.

### 4.1. LLMs for Financial Tasks

#### 4.1.1. General-Purpose LLMs

The rapid advancements in LLMs have ushered in a new era of innovation across various sectors, with finance being a significant beneficiary. Over the past few years, general-purpose LLMs such as GPT-4 have been extensively studied by researchers. Teixeira et al. [31] and Krause [32] presented comprehensive guides to prompt usage in LLMs for financial analysis. Rane et al. [33] explored a comparative analysis of Gemini and ChatGPT, focusing on the discussion of these models' effectiveness and performance in finance and accounting tasks. LLMs have also been tested on a wide range of financial text analytics tasks, demonstrating their versatility and effectiveness [13–17,34,35].

General-purpose LLMs can also be used as investment advisors. the potential of LLMs as financial robo-advisors has been rigorously assessed, generally showing good performance [36–39]. Furthermore, Lu et al. [40] demonstrated ChatGPT could potentially generate portfolios that outperform the markets in out-of-sample tests. Additionally, an innovative LLM multi-agent framework endowed with layered memories has been proposed for stock and fund trading [41]. Their capabilities extend to financial decision-making [42,43], financial auditing [44], financial regulatory interpretation [45], financial budgeting [46], financial risk management [47,48], and analyzing climate change issues related to finance [49].

The reasoning capabilities of LLMs have also been studied extensively. Yu et al. [50,51] examined LLMs' ability for explainable financial time series forecasting, demonstrating these models can generate well-reasoned decisions. Srivastava et al. [52] explored the mathematical reasoning abilities of LLMs on tabular question-answering datasets. Additionally, a comparative study of LLMs for personal financial decision-making in a low resource language, Yoruba language, was conducted by Sikiru et al. [53]. The results indicate that the performance of LLMs is poor compared to their performance with English financial data, highlighting the need for improvements in low-resource languages. To address the high GPU memory consumption associated with LLMs, Liu et al. [54] presented high-performance GPU-based methods for pretraining and fine-tuning LLMs for financial applications.

#### 4.1.2. Finance-Specific LLMs

General-purpose LLMs offer versatility and adaptability for a wide range of financial tasks but may lack the specialized domain knowledge required for complex financial analyses. In contrast, finance-specific LLMs are exclusively trained on financial data. For instance, BloombergGPT, trained on a diverse range of financial data, showcased superior performance in financial tasks compared to existing general-purpose LLMs [55]. Similarly, FinMA was introduced by fine-tuning LLaMA with a tailored dataset, enabling it to execute various financial tasks [56]. Furthermore, FinGPT emerged as an open-source LLM tailored for the finance sector, providing accessible and transparent resources for

researchers and practitioners to develop their FinLLMs [18,19]. Additionally, Li et al. [57] proposed a financial LLM (CFLLM) specially designed to adeptly manage financial texts.

On the other hand, LLMs fine-tuned by financial datasets leverage general-purpose models and enhance their performance in finance-rated tasks through specialized training. For example, Zhang et al. [58] proposed a simple instruction tuning approach to fine-tune general-purpose LLMs, achieving remarkable achievements in financial sentiment analysis. Similarly, Yin et al. [20] introduced FinPT, fine-tuned on LLMs with natural-language customer profile text for predictive purposes and pre-trained by a dataset containing Chinese financial data, and a general-purpose dataset. Additionally, Yang et al. [59] presented, InvestLM, a financial LLM for investment, tuned on LLaMA with a financial investment instruction dataset. Chen et al. [60] proposed a multiple-experts fine-tuning framework to DISC-FinLLM, a large Chinese financial LLM. Finally, Chu et al. [61] created FLLM, a financial LLM employing multitask prompt-based fine-tuning for data pre-processing and pre-understanding, employing abductive augmentation reasoning (AAR) to overcome manual annotation costs.

Multimodal financial LLMs combine the power of language understanding with the rich information contained in financial data across multiple modalities. By integrating textual, numerical, and visual data, multimodal financial LLMs offer a holistic understanding of financial information, enabling more accurate analyses, predictions, and decision-making in the financial domain. For instance, Wang et al. [62] introduced FinVis-GPT, a pioneering multimodal LLM designed to interpret financial charts, marking a significant advancement in the application of multimodal LLMs in finance. Similarly, Bhatia et al. [63] proposed a multimode financial LLM that integrates textual, numerical, tabular, and image financial data, surpassing the performance of ChatGPT-3.5 in financial tasks.

By leveraging advanced language understanding capabilities and domain-specific knowledge, non-English financial LLMs enable more accurate and nuanced analyses of financial information in languages such as Japanese, Spanish, and beyond. For instance, Hirano [64] developed a Japanese financial-specific LLM through continual pre-training, while Zhang et al. [65] introduced FinMA-ES, an LLM tailored for bilingual financial applications aimed at bridging the gap between Spanish and English financial natural language processing (NLP) capabilities.

#### 4.1.3. Benchmarks of LLMs in Finance

Benchmarks are crucial in evaluating the performance of LLMs in the financial domain. These benchmarks serve as standardized tests or datasets against which various LLMs are assessed, allowing researchers, practitioners, and developers to compare and contrast the effectiveness of different models in handling financial tasks. For instance, Xie et al. [56] proposed a standardized benchmark covering a range of financial tasks, while Zhang et al. [66] introduced FinEval, a benchmark specially tailored for the financial domain knowledge in LLMs, including multiple-choice questions covering various topics in business. Guo et al. [67] presented FinLMEval, offering a comprehensive evaluation of LLMs in financial NLP, and Xie et al. [68] proposed an open-sourced evaluation benchmark encompassing 35 datasets across 23 financial tasks.

Additionally, Yin et al. [20] provided high-quality datasets on financial risks, Lei et al. [69] introduced CFBenchmark for evaluating LLMs for financial assistance and Islam et al. [70] proposed FinanceBench for assessing LLMs's performance on open book financial question answering (QA). Zhang et al. [71] curated a practical Text-to-SQL benchmark dataset, beneficial for financial professionals less-skilled in SQL programming. Li et al. [72] introduced AlphaFin dataset for pretraining or fine-tuning financial analysis LLMs, combining traditional research datasets, real-time financial datasets, and handwritten chain-of-thought (COT) data. Furthermore, Hirano [73] constructed a benchmark specific to the Japanese financial domain, while Xu et al. [74] proposed a benchmark for evaluating Chinese-native financial LLMs.



## 4.2. The Risk and Challenge of Generative AI

### 4.2.1. Hallucination

In the realm of GAI, the phenomenon of hallucination manifests when a model generates data or information that deviates from accurately representing the underlying patterns or realities. This occurrence can engender misleading predictions, instill false impressions, or prompt erroneous conclusions based on the generated data. The ramifications of such hallucination are particularly acute within the financial domain, where precision in data and prognostications holds paramount importance for informed decision-making processes. An illustrative case in point is the investigation into GPT-3's efficacy in analyzing climate change related to its financial implications, as undertaken by Leippold [49]. Nevertheless, the inquiry unveiled an issue of hallucination during an interview with the GPT-3 model. Further empirical examination the financial tasks' hallucination behaviors was conducted by Kang and Liu [75], who evaluated the efficacy of various methods, such as few-shot learning and the retrieval augmentation generation method, in mitigating hallucination in LLMs. Their findings underscored the substantial presence of hallucination behaviors in off-the-shelf LLMs when applied to financial tasks.

Consequently, it is imperative that forthcoming research endeavors prioritize strategies aimed at circumventing hallucination by GAI models. Roychowdhury [76] delineated three major stages to design hallucination-minimized LLM-based solutions tailored for the financial domain's decision-makers: prototyping, scaling, and LLM evolution using feedback. These measures ensure that GAI chatbots, autonomous reports, and alerts are reliable and high-quality, thereby facilitating key decision-making processes.

### 4.2.2. Ethical and Social Impact

Generative artificial intelligence (GAI) represents a remarkable tool for discerning users, yet it necessitates critical reflection on the ethical implications and societal ramifications of its integration into the financial industry. Rane [21] extensively explored the multifaceted role and challenges encountered by GAI tools within the intricate realms of finance and accounting, elucidating both their transformative potential and the hurdles that must be overcome to genuinely revolutionize the financial landscape. Similarly, Kalia [77] scrutinized the impact of GAI on the financial sector, advocating for the establishment of robust privacy frameworks, stringent enforcement of data protection regulations, and the promotion of responsible and ethical deployment of these technologies by organizations and policymakers. Sarker [78] delved into the myriad perspectives on LLMs, highlighting both their potentiality and associated risk factors, underscored by heightened awareness. Additionally, Krause [79] deliberated on the potential risks posed by GAI tools in finance and proposed comprehensive mitigation strategies for businesses, emphasizing the importance of employing GAI tools within closed networks, utilizing secure training data, the implementation of robust security measures, providing employee training, and outputs' monitor.

Moreover, Remolina [80] emphasized the necessity of context and sector-specific debates to effectively address the risks and challenges inherent in GAI deployment. For instance, generating financial advice content entails different societal implications than generating imagery of a turtle. Lo and Ross [81] focused on three primary challenges confronting most LLM applications: domain-specific expertise tailored to users' unique circumstances, adherence to moral and ethical standards, and compliance with regulatory guidelines and oversight. Lastly, Yusof and Roslan [82] underscored the imperative of continuous evaluation and adaptation of AI technologies within banking to simultaneously maximize benefits and mitigate associated risks. Collectively, these perspectives underscore the intricate interplay between technological advancement, ethical considerations, and regulatory imperatives within the financial domain, urging stakeholders to navigate these complexities with foresight and diligence.

#### 4.2.3. Financial Regulation

The emergence of robo-advisors and advanced GAI models such as GPT-4 and ChatGPT heralds efficiency gains but also presents distinctive regulatory hurdles. Caspi et al. [83] delved into the regulatory landscape surrounding financial advice in an era increasingly shaped by GAI. Their study scrutinized the extant legal framework governing investment advisors and broker-dealers in the United States, while also examining the ascendancy and impact of robo-advisors and GAI. The pivotal role assumed by AI in the provision of financial advice has necessitated a judicious approach to regulatory strategies. Each regulatory strategy, whether predicated on disclosure mandates or outright prohibitions, carries its own array of benefits and potential challenges. A robust regulatory framework demands more than a cursory understanding; it mandates a comprehensive grasp of these AI technologies and their ramifications.

#### 4.3. *Synthetic Financial Data Generation*

##### 4.3.1. Challenges of Generating Synthetic Data

The financial services sector produces an enormous amount of highly intricate and diverse data. This data is frequently compartmentalized within organizations for several reasons, such as regulatory compliance and operational requirements. Consequently, the sharing of data both across various business units and with external entities like the research community is greatly restricted. Therefore, exploring techniques for creating synthetic financial datasets that maintain the characteristics of real data while ensuring the privacy of the involved parties is crucial. Assefa et al. [84] emphasized the growing need for the financial domain's effective synthetic data generation and highlighted the following academic community's three areas of focus:

- Realistic synthetic datasets generation
- Similarities calculation between real and generated datasets
- Privacy constraints ensuring of the generative process

While these challenges are also present in other domains, the financial sector's additional regulatory and privacy requirements add a layer of complexity. This presents a unique opportunity to study synthetic data generation within the context of financial services.

##### 4.3.2. Existing Works by VAE, GAN, and Diffusion Models

From the literature, it is evident that most existing works on financial synthetic data generation employ Generative Adversarial Networks (GANs). For instance, Zhang et al. [85] introduced a novel GAN architecture to forecast stock closing prices, utilizing a Long Short-Term Memory (LSTM) network as the generator. The LSTM generator captures the data distributions of stocks from the given market data, generating data with similar distributions. Takahashi et al. [86] developed the FIN-GAN model for financial time-series modeling, which learns the properties of data and generates realistic time-series data.

Koshiyama et al. [87] proposed using conditional GANs (cGANs) for trading strategy calibration and aggregation, utilizing the generated samples for ensemble modeling. Bezzina [88] examined the correlation characteristics of synthetic financial time series data generated by TimeGAN, demonstrating that TimeGAN preserves the correlation structure in multi-stock datasets. Ramzan et al. [89] explored GANs for generating synthetic data, emphasizing the generation of datasets that mimic the statistical properties of input data without revealing sensitive information. Vuletić et al. [90] investigated GANs for probability forecasting of financial time series, using a novel economics-driven loss function in the generator. Ljung [91] assessed CTGAN's ability to generate synthetic data, while He and Kita [92] employed a hybrid sequential GANs model with three training strategies using S&P500 data.

In addition to GANs, Variational Autoencoders (VAEs) and Diffusion models have also been employed for financial synthetic data generation. Dalmaso et al. [93] introduced PayVAE, a generative

model designed to learn the temporal and relational structure of financial transactions directly from data. Applied to a real peer-to-peer payments dataset, PayVAE demonstrated its capability to generate realistic transactions. Huang et al. [94] developed a novel generative framework called FTS-Diffusion, which consists of three modules designed to model irregular and scale-invariant patterns in financial time series.

## 5. Conclusions

This research unveils significant theoretical and managerial implications, crucial for comprehending and harnessing the potential of advanced GAI technologies within the financial domain. By summarizing key past research themes, the paper elucidates the evolving landscape of AI technologies and their applications in finance, providing a comprehensive synthesis that advances understanding and informs future research directions.

### 5.1. Theoretical Contribution

Theoretically, this review highlights the transformative potential of LLMs within the financial domain. Building upon previous research that summarises LLMs in finance [23], it emphasizes the need to differentiate between general-purpose and finance-specific LLMs by categorizing research based on training data and application areas. The review then synthesizes findings on the performance and capabilities of prominent LLMs, such as GPT-4 and BloombergGPT, in diverse financial tasks encompassing text analysis, investment advisory, and decision support. This synthesis establishes a foundational framework for future research endeavors to explore critical aspects like performance benchmarks, evaluation criteria, and optimization strategies tailored for LLMs operating in financial contexts. Additionally, the discussion on LLMs' reasoning abilities and their application in financial forecasting and decision-making underscores crucial areas for theoretical exploration. These areas include the development of models capable of generating well-reasoned financial decisions and the enhancement of LLMs to operate effectively within low-resource language environments.

Furthermore, the review delves into the significant ethical and risk considerations surrounding GAI models within the financial sector. Previous research has highlighted the need to address ethical, risk, and synthetic data considerations of GAI [21]. By critically examining the phenomenon of hallucination [76], the paper contributes to the theoretical comprehension of the risks associated with GAI. Additionally, the exploration of ethical concerns, encompassing data privacy and responsible AI utilization, lays the groundwork for the development of robust ethical guidelines and frameworks [22,95]. These frameworks can serve as a foundation for guiding the deployment of GAI technologies in a manner that adheres to ethical principles and aligns with societal expectations. This theoretical foundation is paramount in ensuring that future research and applications of AI in finance are demonstrably ethical and in accordance with societal norms.

The review further underscores the critical role of synthetic financial data generation through the utilization of advanced models like GANs, VAEs, and Diffusion models. By synthesizing past research on methodologies for constructing realistic synthetic datasets while adhering to privacy concerns and regulatory compliance, the paper contributes to the theoretical underpinnings of synthetic data generation in a financial context. This synthesis serves as a roadmap for future research endeavors to refine the accuracy, utility, and ethical considerations associated with synthetic financial data.

### 5.2. Managerial Implications

This review offers valuable managerial insights into the practical applications of GAI technologies within the financial sector. By comprehensively summarizing the capabilities of LLMs across various financial tasks, the paper serves as a practical guide for managers seeking to leverage these models to optimize operational efficiency and enhance decision-making processes. The analysis of performance

metrics and evaluation criteria for LLMs equips managers with the tools necessary to develop effective implementation strategies for integrating these technologies into financial operations [23].

Furthermore, the review emphasizes the critical need to address ethical considerations and potential risks associated with the deployment of GAI technologies. Financial managers can benefit from the comprehensive frameworks and risk management strategies outlined within the review, which provide guidance for the responsible use of AI and the mitigation of potential legal and reputational risks. This proactive approach ensures the ethical and responsible implementation of AI technologies, fostering trust and confidence amongst stakeholders [45].

Finally, the review highlights the practical applications of synthetic financial data generation, offering managers valuable insights into how these technologies can be leveraged to enhance decision-making and bolster the resilience of financial institutions. By investing in the generation and utilization of synthetic financial data, managers gain access to robust datasets that can be employed for financial modeling, stress testing, and scenario analysis, ultimately leading to improved strategic planning and risk management capabilities [84].

5.3. Future Research Agenda

Our research advocates for a future research agenda that prioritizes the exploration of a synergistic relationship between three key areas: differentiation and performance, ethical and risk considerations, and synthetic financial data generation. As depicted in Figure 4, these domains interact dynamically, where advancements in one area can amplify progress in the others.

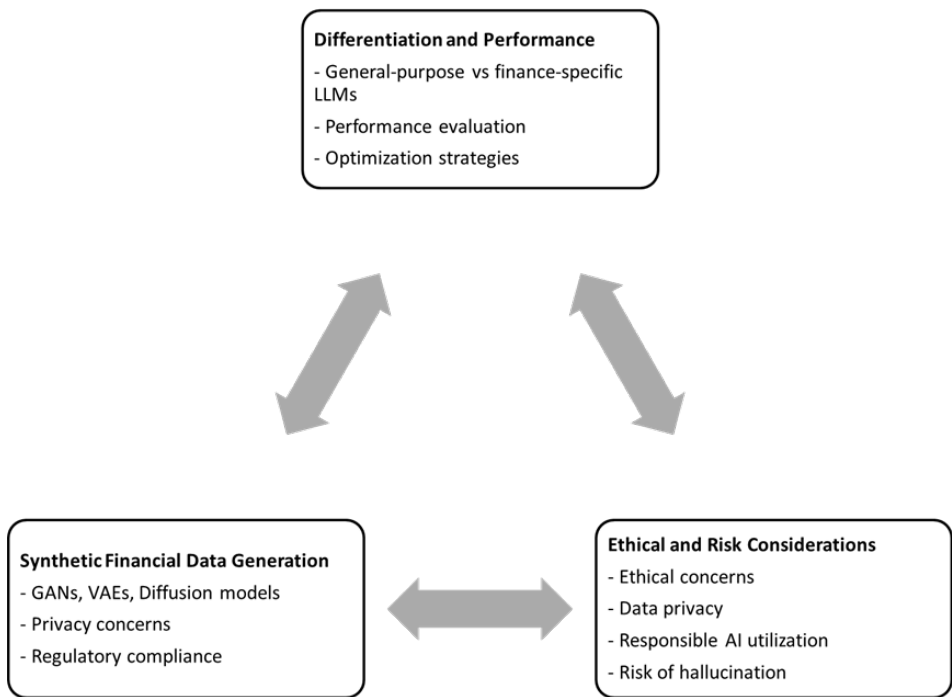


Figure 4. Future Research Agenda.

5.3.1. Intertwined Ethics and Performance Optimization

Central to this approach is the integration of ethical considerations within the performance optimization framework for LLMs employed in finance. This entails ensuring that while LLMs are optimized for financial tasks, the resultant outcomes remain fair, transparent, and accountable. Ethical considerations encompassing data privacy, responsible utilization of AI, and the potential for model hallucination (generating irrelevant or misleading outputs) must be embedded within performance evaluation and optimization strategies. This ensures that advancements in model performance do

not come at the expense of ethical standards, fostering the development of trustworthy and socially responsible AI applications in the financial domain.

### 5.3.2. Synthetic Data: A Boon for Performance Benchmarking

Synthetic financial data has the potential to serve as a valuable tool for evaluating and benchmarking the performance of LLMs specifically designed for finance, while simultaneously safeguarding privacy concerns. By leveraging advanced techniques like Generative Adversarial Networks (GANs), Variational Autoencoders (VAEs), and diffusion models, researchers can create realistic yet privacy-preserving datasets. These datasets can facilitate extensive testing and validation of models, guaranteeing that performance metrics are reliable and reflect real-world scenarios. It is imperative, however, to address privacy concerns and uphold regulatory compliance during synthetic data generation to maintain public trust and adhere to legal requirements.

### 5.3.3. Ethical Considerations in Synthetic Data Generation

A crucial aspect of this research agenda involves addressing the ethical risks associated with synthetic data generation. Researchers must meticulously consider the ethical implications and potential risks to ensure responsible use of this data. This entails guaranteeing that the synthetic data is generated in a manner that does not inadvertently introduce biases or other ethical concerns. Specific considerations include potential biases in data generation, the possibility of data misuse, and the need for robust regulatory frameworks to guide ethical practices. By adhering to these principles, researchers can uphold the integrity and reliability of the models trained on such data.

By pursuing research that investigates these interconnected areas, significant contributions can be made towards the development of robust, ethical, and high-performing LLMs within the financial sector. This holistic approach fosters a collaborative environment where advancements in one area bolster developments in the others, ultimately leading to a more integrated and responsible application of LLMs in the realm of finance. Figure X serves as a visual representation of the proposed future research agenda, highlighting the critical areas of focus and their interconnections, thereby providing a roadmap for researchers to create comprehensive and ethically sound financial AI systems.

In conclusion, the integration of generative AI into the financial sector is revolutionizing data analysis, predictive modeling, and synthetic data generation. This review highlights the significant advancements made with key models like VAEs, GANs, LLMs, and Diffusion models. While these technologies offer substantial benefits, challenges such as data privacy, model interpretability, and regulatory compliance must be addressed. Future research should focus on enhancing model robustness, transparency, and ethical standards to ensure the responsible and effective deployment of generative AI in finance. By tackling these challenges, the financial industry can fully realize the transformative potential of generative AI.

**Author Contributions:** The conceptualization, methodology, validation, formal analysis, investigation, and writing of the paper were completed by all authors equally. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research is funded by MOE Start-up Research Funding (RFE23003).

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Kingma, D.P.; Welling, M. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114* **2013**.
2. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial nets. *Advances in neural information processing systems* **2014**, 27.
3. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. *Advances in neural information processing systems* **2017**, 30.



4. Radford, A.; Narasimhan, K.; Salimans, T.; Sutskever, I.; others. Improving language understanding by generative pre-training **2018**.
5. Radford, A.; Wu, J.; Child, R.; Luan, D.; Amodei, D.; Sutskever, I.; others. Language models are unsupervised multitask learners. *OpenAI blog* **2019**, 1, 9.
6. Brown, T.; Mann, B.; Ryder, N.; Subbiah, M.; Kaplan, J.D.; Dhariwal, P.; Neelakantan, A.; Shyam, P.; Sastry, G.; Askell, A.; others. Language models are few-shot learners. *Advances in neural information processing systems* **2020**, 33, 1877–1901.
7. Ouyang, L.; Wu, J.; Jiang, X.; Almeida, D.; Wainwright, C.; Mishkin, P.; Zhang, C.; Agarwal, S.; Slama, K.; Ray, A.; others. Training language models to follow instructions with human feedback. *Advances in neural information processing systems* **2022**, 35, 27730–27744.
8. Devlin, J.; Chang, M.W.; Lee, K.; Toutanova, K. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805* **2018**.
9. Ho, J.; Jain, A.; Abbeel, P. Denoising diffusion probabilistic models. *Advances in neural information processing systems* **2020**, 33, 6840–6851.
10. Bollen, J.; Mao, H.; Zeng, X. Twitter mood predicts the stock market. *Journal of computational science* **2011**, 2, 1–8.
11. Wisniewski, T.P.; Yekini, L.S. Stock market returns and the content of annual report narratives. *Accounting Forum*. Elsevier, 2015, Vol. 39, pp. 281–294.
12. McGurk, Z.; Nowak, A.; Hall, J.C. Stock returns and investor sentiment: textual analysis and social media. *Journal of Economics and Finance* **2020**, 44, 458–485.
13. Gupta, U. Gpt-investar: Enhancing stock investment strategies through annual report analysis with large language models. *arXiv preprint arXiv:2309.03079* **2023**.
14. Fatemi, S.; Hu, Y. A Comparative Analysis of Fine-Tuned LLMs and Few-Shot Learning of LLMs for Financial Sentiment Analysis. *arXiv preprint arXiv:2312.08725* **2023**.
15. Li, X.; Chan, S.; Zhu, X.; Pei, Y.; Ma, Z.; Liu, X.; Shah, S. Are ChatGPT and GPT-4 General-Purpose Solvers for Financial Text Analytics? A Study on Several Typical Tasks. *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing: Industry Track, 2023*, pp. 408–422.
16. Pavlyshenko, B.M. Financial News Analytics Using Fine-Tuned Llama 2 GPT Model. *arXiv preprint arXiv:2308.13032* **2023**.
17. Xing, F. Designing Heterogeneous LLM Agents for Financial Sentiment Analysis. *arXiv preprint arXiv:2401.05799* **2024**.
18. Liu, X.Y.; Wang, G.; Zha, D. Fingpt: Democratizing internet-scale data for financial large language models. *arXiv preprint arXiv:2307.10485* **2023**.
19. Yang, H.; Liu, X.Y.; Wang, C.D. Fingpt: Open-source financial large language models. *arXiv preprint arXiv:2306.06031* **2023**.
20. Yin, Y.; Yang, Y.; Yang, J.; Liu, Q. FinPT: Financial Risk Prediction with Profile Tuning on Pretrained Foundation Models. *arXiv preprint arXiv:2308.00065* **2023**.
21. Rane, N. Role and Challenges of ChatGPT and Similar Generative Artificial Intelligence in Finance and Accounting. *Available at SSRN 4603206* **2023**.
22. Ding, Q.; Ding, D.; Wang, Y.; Guan, C.; Ding, B. Unraveling the landscape of large language models: a systematic review and future perspectives. *Journal of Electronic Business & Digital Economics* **2023**.
23. Li, Y.; Wang, S.; Ding, H.; Chen, H. Large language models in finance: A survey. *Proceedings of the Fourth ACM International Conference on AI in Finance, 2023*, pp. 374–382.
24. Lee, J.; Stevens, N.; Han, S.C.; Song, M. A Survey of Large Language Models in Finance (FinLLMs). *arXiv preprint arXiv:2402.02315* **2024**.
25. Barde, K.; Kulkarni, P.A. Applications of Generative AI in Fintech. *Proceedings of the Third International Conference on AI-ML Systems, 2023*, pp. 1–5.
26. Krause, D. Large Language Models and Generative AI in Finance: An Analysis of ChatGPT, Bard, and Bing AI. *Bard, and Bing AI (July 15, 2023)* **2023**.
27. Mbanyele, W. Generative AI and ChatGPT in Financial Markets and Corporate Policy: A Comprehensive Review. *Available at SSRN 4745990* **2024**.
28. Grootendorst, M. BERTopic: Neural topic modeling with a class-based TF-IDF procedure. *arXiv preprint arXiv:2203.05794* **2022**.

29. Blei, D.M.; Ng, A.Y.; Jordan, M.I. Latent dirichlet allocation. *Journal of machine Learning research* **2003**, *3*, 993–1022.
30. Févotte, C.; Idier, J. Algorithms for nonnegative matrix factorization with the  $\beta$ -divergence. *Neural computation* **2011**, *23*, 2421–2456.
31. Teixeira, A.C.; Marar, V.; Yazdanpanah, H.; Pezente, A.; Ghassemi, M. Enhancing Credit Risk Reports Generation using LLMs: An Integration of Bayesian Networks and Labeled Guide Prompting. *Proceedings of the Fourth ACM International Conference on AI in Finance, 2023*, pp. 340–348.
32. Krause, D. Proper generative ai prompting for financial analysis. *Available at SSRN 4453664* **2023**.
33. Rane, N.; Choudhary, S.; Rane, J. Gemini or ChatGPT? Efficiency, Performance, and Adaptability of Cutting-Edge Generative Artificial Intelligence (AI) in Finance and Accounting. *Efficiency, Performance, and Adaptability of Cutting-Edge Generative Artificial Intelligence (AI) in Finance and Accounting (February 19, 2024)* **2024**.
34. Callanan, E.; Mbakwe, A.; Papadimitriou, A.; Pei, Y.; Sibue, M.; Zhu, X.; Ma, Z.; Liu, X.; Shah, S. Can gpt models be financial analysts? an evaluation of chatgpt and gpt-4 on mock cfa exams. *arXiv preprint arXiv:2310.08678* **2023**.
35. Zhao, H.; Liu, Z.; Wu, Z.; Li, Y.; Yang, T.; Shu, P.; Xu, S.; Dai, H.; Zhao, L.; Mai, G.; others. Revolutionizing finance with llms: An overview of applications and insights. *arXiv preprint arXiv:2401.11641* **2024**.
36. Niszczota, P.; Abbas, S. GPT has become financially literate: Insights from financial literacy tests of GPT and a preliminary test of how people use it as a source of advice. *Finance Research Letters* **2023**, *58*, 104333.
37. Lakkaraju, K.; Vuruma, S.K.R.; Pallagani, V.; Muppasani, B.; Srivastava, B. Can llms be good financial advisors?: An initial study in personal decision making for optimized outcomes. *arXiv preprint arXiv:2307.07422* **2023**.
38. Huang, Z.; Che, C.; Zheng, H.; Li, C. Research on Generative Artificial Intelligence for Virtual Financial Robo-Advisor. *Academic Journal of Science and Technology* **2024**, *10*, 74–80.
39. Liu, Y.; Wang, J. Analysis of Financial Market using Generative Artificial Intelligence. *Academic Journal of Science and Technology* **2024**, *11*, 21–25.
40. Lu, F.; Huang, L.; Li, S. ChatGPT, Generative AI, and Investment Advisory. *Available at SSRN* **2023**.
41. Li, Y.; Yu, Y.; Li, H.; Chen, Z.; Khashanah, K. Tradinggpt: Multi-agent system with layered memory and distinct characters for enhanced financial trading performance. *arXiv preprint arXiv:2309.03736* **2023**.
42. Yu, Y.; Li, H.; Chen, Z.; Jiang, Y.; Li, Y.; Zhang, D.; Liu, R.; Suchow, J.W.; Khashanah, K. FinMem: A performance-enhanced LLM trading agent with layered memory and character design. *Proceedings of the AAAI Symposium Series, 2024, Vol. 3*, pp. 595–597.
43. Lakkaraju, K.; Jones, S.E.; Vuruma, S.K.R.; Pallagani, V.; Muppasani, B.C.; Srivastava, B. LLMs for Financial Advisement: A Fairness and Efficacy Study in Personal Decision Making. *Proceedings of the Fourth ACM International Conference on AI in Finance, 2023*, pp. 100–107.
44. Hillebrand, L.; Berger, A.; Deußer, T.; Dilmaghani, T.; Khaled, M.; Kliem, B.; Loitz, R.; Pielka, M.; Leonhard, D.; Bauckhage, C.; others. Improving zero-shot text matching for financial auditing with large language models. *Proceedings of the ACM Symposium on Document Engineering 2023, 2023*, pp. 1–4.
45. Cao, Z.; Feinstein, Z. Large Language Model in Financial Regulatory Interpretation. *arXiv preprint arXiv:2405.06808* **2024**.
46. de Zarzà, I.; de Curtò, J.; Roig, G.; Calafate, C.T. Optimized financial planning: Integrating individual and cooperative budgeting models with llm recommendations. *AI* **2023**, *5*, 91–114.
47. Chen, B.; Wu, Z.; Zhao, R. From fiction to fact: the growing role of generative AI in business and finance. *Journal of Chinese Economic and Business Studies* **2023**, *21*, 471–496.
48. Wang, Y. Generative AI in Operational Risk Management: Harnessing the Future of Finance. *Operational Risk Management: Harnessing the Future of Finance (May 17, 2023)* **2023**.
49. Leippold, M. Thus spoke GPT-3: Interviewing a large-language model on climate finance. *Finance Research Letters* **2023**, *53*, 103617.
50. Yu, X.; Chen, Z.; Lu, Y. Harnessing LLMs for temporal data-a study on explainable financial time series forecasting. *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing: Industry Track, 2023*, pp. 739–753.
51. Yu, X.; Chen, Z.; Ling, Y.; Dong, S.; Liu, Z.; Lu, Y. Temporal Data Meets LLM-Explainable Financial Time Series Forecasting. *arXiv preprint arXiv:2306.11025* **2023**.

52. Srivastava, P.; Malik, M.; Ganu, T. Assessing LLMs' Mathematical Reasoning in Financial Document Question Answering. *arXiv preprint arXiv:2402.11194* **2024**.
53. Sikiru, R.D.; Adekanmbi, O.; Soronnadi, A. Comparative Study of LLMs for Personal Financial Decision in Low Resource Language. 5th Workshop on African Natural Language Processing.
54. Liu, X.Y.; Zhang, J.; Wang, G.; Tong, W.; Walid, A. FinGPT-HPC: Efficient Pretraining and Finetuning Large Language Models for Financial Applications with High-Performance Computing. *arXiv preprint arXiv:2402.13533* **2024**.
55. Wu, S.; Irsoy, O.; Lu, S.; Dabrovolski, V.; Dredze, M.; Gehrmann, S.; Kambadur, P.; Rosenberg, D.; Mann, G. BloombergGPT: A large language model for finance, 2023. *ArXiv preprint: <https://arxiv.org/pdf/2303.17564.pdf>* **2024**.
56. Xie, Q.; Han, W.; Zhang, X.; Lai, Y.; Peng, M.; Lopez-Lira, A.; Huang, J. Pixiu: A large language model, instruction data and evaluation benchmark for finance. *arXiv preprint arXiv:2306.05443* **2023**.
57. Li, J.; Bian, Y.; Wang, G.; Lei, Y.; Cheng, D.; Ding, Z.; Jiang, C. Cfgpt: Chinese financial assistant with large language model. *arXiv preprint arXiv:2309.10654* **2023**.
58. Zhang, B.; Yang, H.; Liu, X.Y. Instruct-fingpt: Financial sentiment analysis by instruction tuning of general-purpose large language models. *arXiv preprint arXiv:2306.12659* **2023**.
59. Yang, Y.; Tang, Y.; Tam, K.Y. Investlm: A large language model for investment using financial domain instruction tuning. *arXiv preprint arXiv:2309.13064* **2023**.
60. Chen, W.; Wang, Q.; Long, Z.; Zhang, X.; Lu, Z.; Li, B.; Wang, S.; Xu, J.; Bai, X.; Huang, X.; others. Disc-finllm: A chinese financial large language model based on multiple experts fine-tuning. *arXiv preprint arXiv:2310.15205* **2023**.
61. Chu, Z.; Guo, H.; Zhou, X.; Wang, Y.; Yu, F.; Chen, H.; Xu, W.; Lu, X.; Cui, Q.; Li, L.; others. Data-centric financial large language models. *arXiv preprint arXiv:2310.17784* **2023**.
62. Wang, Z.; Li, Y.; Wu, J.; Soon, J.; Zhang, X. Finvis-gpt: A multimodal large language model for financial chart analysis. *arXiv preprint arXiv:2308.01430* **2023**.
63. Bhatia, G.; Nagoudi, E.M.B.; Cavusoglu, H.; Abdul-Mageed, M. FinTral: A Family of GPT-4 Level Multimodal Financial Large Language Models. *arXiv preprint arXiv:2402.10986* **2024**.
64. Hirano, M.; Imajo, K. Construction of Domain-specified Japanese Large Language Model for Finance through Continual Pre-training. *arXiv preprint arXiv:2404.10555* **2024**.
65. Zhang, X.; Xiang, R.; Yuan, C.; Feng, D.; Han, W.; Lopez-Lira, A.; Liu, X.Y.; Ananiadou, S.; Peng, M.; Huang, J.; others. D'olares or Dollars? Unraveling the Bilingual Prowess of Financial LLMs Between Spanish and English. *arXiv preprint arXiv:2402.07405* **2024**.
66. Zhang, L.; Cai, W.; Liu, Z.; Yang, Z.; Dai, W.; Liao, Y.; Qin, Q.; Li, Y.; Liu, X.; Liu, Z.; others. Fineval: A chinese financial domain knowledge evaluation benchmark for large language models. *arXiv preprint arXiv:2308.09975* **2023**.
67. Guo, Y.; Xu, Z.; Yang, Y. Is chatgpt a financial expert? evaluating language models on financial natural language processing. *arXiv preprint arXiv:2310.12664* **2023**.
68. Xie, Q.; Han, W.; Chen, Z.; Xiang, R.; Zhang, X.; He, Y.; Xiao, M.; Li, D.; Dai, Y.; Feng, D.; others. The FinBen: An Holistic Financial Benchmark for Large Language Models. *arXiv preprint arXiv:2402.12659* **2024**.
69. Lei, Y.; Li, J.; Jiang, M.; Hu, J.; Cheng, D.; Ding, Z.; Jiang, C. Cfbenchmark: Chinese financial assistant benchmark for large language model. *arXiv preprint arXiv:2311.05812* **2023**.
70. Islam, P.; Kannappan, A.; Kiela, D.; Qian, R.; Scherrer, N.; Vidgen, B. Financebench: A new benchmark for financial question answering. *arXiv preprint arXiv:2311.11944* **2023**.
71. Zhang, C.; Mao, Y.; Fan, Y.; Mi, Y.; Gao, Y.; Chen, L.; Lou, D.; Lin, J. FinSQL: Model-Agnostic LLMs-based Text-to-SQL Framework for Financial Analysis. *arXiv preprint arXiv:2401.10506* **2024**.
72. Li, X.; Li, Z.; Shi, C.; Xu, Y.; Du, Q.; Tan, M.; Huang, J.; Lin, W. AlphaFin: Benchmarking Financial Analysis with Retrieval-Augmented Stock-Chain Framework. *arXiv preprint arXiv:2403.12582* **2024**.
73. Hirano, M. Construction of a japanese financial benchmark for large language models. *arXiv preprint arXiv:2403.15062* **2024**.
74. Xu, L.; Zhu, L.; Wu, Y.; Xue, H. SuperCLUE-Fin: Graded Fine-Grained Analysis of Chinese LLMs on Diverse Financial Tasks and Applications. *arXiv preprint arXiv:2404.19063* **2024**.

75. Kang, H.; Liu, X.Y. Deficiency of Large Language Models in Finance: An Empirical Examination of Hallucination. *arXiv preprint arXiv:2311.15548* **2023**.
76. Roychowdhury, S. Journey of hallucination-minimized generative ai solutions for financial decision makers. *Proceedings of the 17th ACM International Conference on Web Search and Data Mining*, 2024, pp. 1180–1181.
77. Kalia, S. Potential Impact of Generative Artificial Intelligence (AI) on the Financial Industry. *International Journal on Cybernetics & Informatics (IJCI)* **2023**, 12, 37.
78. Sarker, I.H. LLM potentiality and awareness: a position paper from the perspective of trustworthy and responsible AI modeling. *Discover Artificial Intelligence* **2024**, 4, 40.
79. Krause, D. Mitigating risks for financial firms using generative AI Tools. *Available at SSRN 4452600* **2023**.
80. Remolina, N. Generative AI in Finance: Risks and Potential Solutions. *Finance: Risks and Potential Solutions* (November 9, 2023). *Singapore Management University School of Law Research Paper Forthcoming*, SMU Centre for AI & Data Governance Research Paper Forthcoming **2023**.
81. Lo, A.W.; Ross, J. Can ChatGPT Plan Your Retirement?: Generative AI and Financial Advice. *Generative AI and Financial Advice* (February 11, 2024) **2024**.
82. Yusof, S.A.B.M.; Roslan, F.A.B.M. The Impact of Generative AI in Enhancing Credit Risk Modeling and Decision-Making in Banking Institutions. *Emerging Trends in Machine Intelligence and Big Data* **2023**, 15, 40–49.
83. Caspi, I.; Felber, S.S.; Gillis, T.B. Generative AI and the Future of Financial Advice Regulation. *GenLaw Center* **2023**.
84. Assefa, S.A.; Dervovic, D.; Mahfouz, M.; Tillman, R.E.; Reddy, P.; Veloso, M. Generating synthetic data in finance: opportunities, challenges and pitfalls. *Proceedings of the First ACM International Conference on AI in Finance*, 2020, pp. 1–8.
85. Zhang, K.; Zhong, G.; Dong, J.; Wang, S.; Wang, Y. Stock market prediction based on generative adversarial network. *Procedia computer science* **2019**, 147, 400–406.
86. Takahashi, S.; Chen, Y.; Tanaka-Ishii, K. Modeling financial time-series with generative adversarial networks. *Physica A: Statistical Mechanics and its Applications* **2019**, 527, 121261.
87. Koshiyama, A.; Firoozye, N.; Treleven, P. Generative adversarial networks for financial trading strategies fine-tuning and combination. *Quantitative Finance* **2021**, 21, 797–813.
88. Bezzina, P. Improving portfolio construction using deep generative machine learning models applying generative models on financial market data. Master's thesis, University of Malta, 2023.
89. Ramzan, F.; Sartori, C.; Consoli, S.; Reforgiato Recupero, D. Generative Adversarial Networks for Synthetic Data Generation in Finance: Evaluating Statistical Similarities and Quality Assessment. *AI* **2024**, 5, 667–685.
90. Vuletić, M.; Prenzel, F.; Cucuringu, M. Fin-gan: Forecasting and classifying financial time series via generative adversarial networks. *Quantitative Finance* **2024**, pp. 1–25.
91. Ljung, M. Synthetic Data Generation for the Financial Industry Using Generative Adversarial Networks, 2021.
92. He, B.; Kita, E. Stock price prediction by using hybrid sequential generative adversarial networks. *2020 International Conference on Data Mining Workshops (ICDMW)*. IEEE, 2020, pp. 341–347.
93. Dalmaso, N.; Tillman, R.E.; Reddy, P.; Veloso, M. Payvae: A generative model for financial transactions. *AAAI 2021 Workshop on Knowledge Discovery from Unstructured Data in Financial Services Workshop*, 2021.
94. Huang, H.; Chen, M.; Qiao, X. Generative Learning for Financial Time Series with Irregular and Scale-Invariant Patterns. *The Twelfth International Conference on Learning Representations*, 2023.
95. Guan, C.; Ding, D.; Gupta, P.; Hung, Y.C.; Jiang, Z. A Systematic Review of Research on ChatGPT: The User Perspective. *Exploring Cyber Criminals and Data Privacy Measures* **2023**, pp. 124–150.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.