



# International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

*(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)*



**Impact Factor: 8.699**

**Volume 14, Issue 8, August 2025**

# A Comprehensive Foundation Models in Financial Intelligence: A Comprehensive Survey of LLMs, Diffusion Models, and Multimodal AI

Dr. R. Gayathri, Dr C Bhuvaneshwari

Associate Professor, Department of Computer Science (PG), Kristu Jayanti (Deemed to be University),  
Bengaluru, India

Assistant Professor, Department of Computer Science, Dr. G. R. Damodaran College of Science (Autonomous),  
Coimbatore, India

**ABSTRACT:** The exponential advancement of artificial intelligence has led to the emergence of powerful foundation models such as Large Language Models (LLMs), diffusion models, and multimodal AI systems which are capable of understanding and working on the various input types such as text, images, videos and tabular data. These technologies provide a significant development in the field of financial intelligence, facilitating automated analysis of financial documents, generation of synthetic financial data, detection of anomalies, and support for real-time decision-making. This paper provides a detailed insights on current developments in LLMs, diffusion models, and multimodal AI with a specific focus on the various applications of financial sector. For every class of models, the underlying architectures, benchmark datasets, learning methodologies and assessment measures are analysed. Furthermore, this paper provides significant domain-specific challenges, including inaccuracies in financial text generation, data privacy concerns, regulatory compliance requirements, and the growing need for model interpretability. By conducting a comprehensive review of contemporary research and emerging developments, this survey identifies prevailing limitations within the field. It systematically highlights critical research gaps that hinder progress in financial AI applications. The study further proposes strategic directions for the development of intelligent, trustworthy, and domain-specific AI systems tailored to the financial sector.

**KEYWORDS:** Large Language Models (LLMs), Diffusion Models; Multimodal AI, Financial Intelligence, Financial Document Analysis; Synthetic Financial Data, Cross-Modal Reasoning, Explainable AI (XAI), FinTech, Financial NLP, AI in Finance, Regulatory Compliance, Decision Support Systems.

## I. INTRODUCTION

### Background and Motivation

With the progression of foundation models that demonstrate remarkable generalization across various tasks and modalities, artificial intelligence has undergone a transformative shift. Especially GPT-4, Claude, and PaLM are few examples of Large Language Models (LLMs) that have shown distinguished performance in understanding, generating, and reasoning with human language. Complementing these are diffusion models, which have revolutionized generative AI—particularly in the creation of images and videos—and are increasingly being applied to text and structured data generation. The integration of these models with multimodal AI systems, capable of simultaneously interpreting text, images, videos, and tabular data, has unlocked new possibilities for intelligent automation across different sectors.

These models have considerable promise in the financial industry, where crucial choices are made based on a variety of data formats, including market charts, earnings reports, regulatory documents, and real-time news. Multimodal AI can integrate multiple inputs for comprehensive decision support, diffusion models can be used to create synthetic financial time series or stress-test scenarios, and LLMs can help comprehend and summarise complex financial texts. Financial technologies (FinTech) that are intelligent, flexible, and explicable are made possible by their combined efforts.



## II. LITERATURE SURVEY

### LLM Applications in Finance

This paper presents a practical survey of Large Language Models (LLMs) in financial applications, focusing on current solutions and implementation guidance. It reviews techniques such as zero/few-shot learning, fine-tuning, and custom model training for financial tasks. A decision framework is proposed to help professionals select appropriate LLM solutions based on constraints like data and compute. The survey categorizes LLM applications in finance into five key areas such as linguistic tasks, sentiment analysis, financial time series analysis, financial reasoning and agent-based modelling. The paper also addresses challenges, offering a roadmap for responsibly adopting LLMs in finance [1].

### LLMs in Finance, Marketing, and SCM

This survey investigates how Large Language Models (LLMs) and AI-driven systems transform three pivotal business domains: Finance, Marketing, and Supply Chain Management (SCM). It examines their roles in improving operational efficiency, enabling strategic insights, and stimulating economic growth. Unlike earlier literature that examines isolated applications, this work offers an integrative analysis across these domains, revealing both individual and synergistic benefits of adopting LLMs. In finance, the survey covers applications in market forecasting, risk evaluation, regulatory compliance, and fraud detection. In marketing, it explores capabilities in hyper-personalized consumer engagement, resource optimization, and conversion enhancement. For SCM, it discusses improvements in demand forecasting and logistics optimization, emphasizing agility and resilience in AI-enhanced systems [2].

### Advancements in Multimodal LLMs

This paper provides a comprehensive review of Multimodal Large Language Models (LLMs) that synthesize diverse data types—text, images, and audio—to enrich machine learning capabilities. The authors trace the evolution from early unimodal architectures to advanced multimodal systems such as BERT, GPT-4, CLIP, DALL-E, and Perceiver. This also describes how these models integrate across modalities by designing cross-modal encoders, attention mechanisms, and fusion strategies. Key application areas include image–text generation, multimodal retrieval, and AI-driven cybersecurity and social media analytics. This highlights persistent challenges: aligning heterogeneous inputs, ensuring robust generalization across domains, mitigating biases, and maintaining computational efficiency. Finally, it outlines future research directions: scalable architectures, more interpretable fusion mechanisms, and real-time multimodal processing frameworks—aiming to bridge the gap between performance and usability in real-world systems [3].

### Interpretable Multimodal Large Language Models

The survey addresses the interpretability and explainability of Multimodal Large Language Models (MLLMs)—a growing class of AI that fuses language models with vision, audio, and video inputs. The authors propose a novel analytical framework structured around three dimensions: Data, Model, and Training & Inference. The interpretability techniques are systematically examined, spanning token-level and embedding-level representations, architectural insights, and transparency strategies during model training and inference. By evaluating a range of methods—such as attention visualization, feature attribution, and modular decomposition—this highlights strengths and limitations in current approaches. The survey draws attention to the importance of transparency and trustworthiness in high-stakes applications, like finance and healthcare. It concludes with future research directions addressing robustness, refined interpretability, and benchmark development for multimodal explainability [4].

### A Survey of Financial AI: Architectures, Advances and Open Challenges

This survey systematically examines recent developments in financial AI across three key dimensions: predictive modeling, decision-making frameworks, and knowledge augmentation using unstructured data. It reviews foundational progress in modeling financial time series, the application of graph-based architectures for capturing market relationships, and hierarchical strategies for portfolio optimization. The study highlights important trade-offs between model complexity and real-world deployment, particularly in high-frequency trading environments. A critical gap remains in bridging academic advances with industrial implementation, and the paper outlines open challenges related to practical applicability and performance enhancement. The survey provides a comprehensive taxonomy and analysis intended to inform both researchers and practitioners on the state-of-the-art and directions for future exploration in computational finance [5].

**A Survey of Large Language Models in Finance (FinLLMs)**

This survey presents a comprehensive overview of Financial Large Language Models (FinLLMs), tracing the evolution from general-purpose pretrained language models to those customized for financial applications. A chronological framework is provided, spanning GPT-series models, open-source LLMs, and specialized financial-domain models. Five key techniques are compared across models, including training methodologies, data sources, fine-tuning strategies, and instruction-based tuning. Performance evaluation is summarized on six benchmark datasets, and eight advanced financial NLP tasks are cataloged to support the development of more capable FinLLMs. The study identifies critical challenges—such as hallucination, data privacy, and computational efficiency—and consolidates datasets and evaluation benchmarks to support further research [6].

**Table: 1 Survey Comparison Table**

Survey	Financial LLMs	Benchmarks	Applications	Challenges
LLM Applications in Finance [1]	✓	✗	✓	✓
LLMs in Finance, Marketing, and SCM [2]	✓	✗	✓	✗
Advancements in Multimodal LLMs [3]	✗	✗	✓	✓
Interpretable Multimodal LLMs [4]	✗	✗	✓	✓
A Survey of Financial AI [5]	✓	✗	✓	✓
A Survey of Large Language Models in Finance (FinLLMs) [6]	✓	✓	✓	✓
Ours	✓	✓	✓	✓

This table summarizes prior literature in terms of five key aspects. The Survey column lists shorthand references (e.g., [1], [2]) corresponding to the papers cited in the literature review, enabling readers to link back to the bibliography without restating full titles. The Financial LLMs column indicates whether the paper specifically focuses on large language models in financial domains such as stock prediction, risk modeling, or fraud detection, where a ✓ denotes inclusion and a ✗ indicates absence of finance-specific focus. The Benchmarks column captures whether the work discusses benchmark datasets, evaluation metrics, or standardized testing frameworks for LLMs, which are essential for replicability and comparative analysis of models. The Applications column identifies whether practical use cases and real-world deployments of LLMs are covered, such as in finance, marketing, supply chain, healthcare, or multimodal settings. Finally, the Challenges column reflects whether the study addresses open problems or limitations in LLMs, including issues like bias, computational cost, explainability, and domain adaptation.

**III. CONTRIBUTIONS****The primary contributions of this survey are summarized as follows:**

Comprehensive Analysis of Foundation Models in Finance – A unified review is presented covering Large Language Models (LLMs), diffusion models, and multimodal AI systems, with emphasis on their applications in the financial sector. The discussion includes architectural insights, training methodologies, benchmark datasets, and evaluation frameworks.

Domain-Specific Applications and Use Cases – Financial applications are systematically categorized and analyzed, including financial document analysis, synthetic data generation, anomaly detection, market forecasting, and decision-support systems.

Comparative Benchmark Study – Benchmark datasets, performance metrics, and evaluation protocols across multiple studies are consolidated to enable reproducibility and facilitate model comparison for financial AI research.

Identification of Critical Challenges – Key limitations are highlighted, such as inaccuracies in financial text generation, data privacy concerns, regulatory compliance requirements, and the necessity for interpretable AI in finance.

Strategic Research Directions – Future research pathways are proposed to address identified gaps, focusing on the development of intelligent, trustworthy, and finance-specific AI systems capable of operating within regulatory and ethical boundaries.

#### **IV. TECHNICAL BACKGROUND ON FOUNDATION MODELS**

Section 2 provides a comprehensive overview of the core architectures underpinning Large Language Models (LLMs), diffusion models, and multimodal AI systems—forming the foundational context for the survey.

**Large Language Models (LLMs):** Foundational models such as GPT, BERT, and domain-tuned variants like FinBERT are introduced, with emphasis on their transformer architectures, self-attention mechanisms, and scalability in pretraining on large financial corpora. Their suitability for tasks requiring contextual comprehension and language generation is highlighted.

**Diffusion Models:** These generative frameworks, originally designed for image synthesis, are examined for their growing relevance in structured time-series generation. Their ability to model complex data distributions makes them promising candidates for synthetic financial data generation and stress-testing scenarios.

**Multimodal AI Systems:** The integration of text, image, video, and structured data via models like CLIP, Perceiver, and emerging multimodal LLMs is detailed. Architectural components such as joint cross-modal embeddings, attention fusion layers, and encoder-decoder schemas are discussed, highlighting the models' ability to process heterogeneous financial data. This section establishes the groundwork required to understand how these models function and interoperate within the financial intelligence domain.

#### **V. LITERATURE SURVEY AND COMPARITHE ANALYSIS**

This section presents a structured literature survey, categorizing existing works based on model type, application scope, and domain relevance, and features a comparative summary table for a clear synthesis.

##### **Financial LLMs:**

Includes surveys detailing zero/few-shot, fine-tuning, and domain-specific LLMs tailored to financial tasks (e.g., textual analysis, sentiment scoring, reasoning, agent-based modeling) [7], as well as open finance models like FinGPT and BloombergGPT which offer domain data-centric approaches and extensive token training [8].

##### **Cross-Domain Surveys:**

Broader integrative literature covering finance alongside marketing and supply chain management contexts helps underscore interdisciplinary applications and shared AI insights in operational efficiency and decision-making [9].

##### **Multimodal Model Reviews:**

Includes technical analyses of multimodal large language models, especially in vision-language tasks, with architectural evaluations as well as discussions of challenges such as model alignment, robustness, and scalability [10]. Interpretability-focused multimodal surveys emphasize transparency approaches including Data/Model/Training frameworks, attention visualization, and modular decomposition [11].

### Comparative Table & Synthesis:

A tabulated comparison highlights each surveyed work in terms of model type, application domain, benchmark coverage, and the challenges addressed—creating a quick reference for researchers. This comparative essay bridges the conceptual gap between architectural advances and their real-world financial deployments.

**Table 2: Benchmarks of LLMs on Financial Applications**

No.	Name / Title	Year	Task	Modality	Model	Language	Open Source
1	FinBERT: A Pre-trained Financial Language Model for Financial Text Mining (IEEE)	2024	Sentiment analysis on financial texts	Text	FinBERT	English	Yes
2	Large Language Models in Finance: Use Cases, Risks, and Opportunities (SSRN)	2024	Financial document analysis, question answering	Text	GPT-based, LLaMA-based models	English	Partial
3	Applying AI Models for Global Financial Data Streams in Web and Grid Services (Inderscience)	2024	Real-time financial data analysis	Text	Custom AI pipeline	English	No
4	Leveraging LLMs for Financial Narrative Processing (HAL)	2024	Financial narrative extraction, event detection	Text	LLaMA, GPT-based	English, French	Partial
5	FinEval: Benchmarking LLMs for Financial Tasks (arXiv)	2024	Financial QA, sentiment, information extraction	Text	Multiple LLMs (GPT, Claude, LLaMA)	English	No
6	AI in Financial Market Forecasting (The Innovation)	2025	Market trend prediction, portfolio optimization	Text, Time-series	Hybrid models (LLMs + statistical models)	English	No
7	FINGPT: Domain-Specific Large Language Models for Finance (IEEE)	2025	Financial document analysis, sentiment, Q&A	Text	FINGPT	English	Yes
8	Evaluating Multimodal LLMs for Financial News and Market Data (arXiv)	2024	Financial news summarization, price prediction	Text, Tabular	Multimodal LLM	English	Partial

## **VI. RESULTS AND DISCUSSION**

From the literature surveyed, it is evident that foundation models have brought significant advancements to financial AI applications. LLMs are powerful in financial text comprehension, reasoning, and sentiment analysis, while diffusion models provide synthetic data generation capabilities useful for financial stress testing and anomaly detection. Multimodal AI, on the other hand, addresses the need for cross-modal integration, enabling decision-making across diverse data formats. However, persistent challenges such as hallucination in financial LLMs, lack of robust interpretability, data privacy issues, and high computational cost remain barriers to large-scale adoption. Comparative studies show that while general-purpose models perform reasonably well, domain-specific models like FinBERT and BloombergGPT achieve superior accuracy in financial tasks. This highlights the importance of domain adaptation. Future research should focus on hybrid approaches, transparent evaluation, regulatory alignment, and computational efficiency to develop trustworthy financial AI systems.

## **VII. CHALLENGES**

Applying foundation models in finance presents notable hurdles. One key issue is the generation of inaccurate or misleading information, where models produce outputs that sound convincing but are factually wrong, which can misguide critical financial decisions. Since most models are trained on general data, adapting them to financial language and domain-specific tasks remains difficult. Synthetic data created by generative models may also lack realism, raising concerns for risk analysis. Additional challenges include combining multiple data types effectively, ensuring explainability for compliance, safeguarding privacy, managing high computational costs, avoiding bias, and integrating these systems into traditional financial infrastructures.

## **VIII. CONCLUSION**

This survey has presented a comprehensive review of foundation models in finance, covering Large Language Models, diffusion models, and multimodal AI. It analyzed benchmark datasets, applications, and challenges while emphasizing the need for domain-specific interpretability and ethical AI practices. The study identified critical limitations and proposed research directions aimed at creating trustworthy and intelligent AI systems tailored to financial applications.

## **REFERENCES**

- [1] Z. Zhang, et al., "Financial Large Language Models: Applications, Techniques, and Opportunities," arXiv preprint arXiv:2406.11903, 2024. [Online]. Available: <https://arxiv.org/pdf/2406.11903>
- [2] A. Gupta, et al., "Large Language Models in Finance, Marketing, and Supply Chain Management," SSRN Electronic Journal, 2024. [Online]. Available: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4932769](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4932769)
- [3] P. Zhang and Y. Chen, "Advancements in Multimodal Large Language Models," in Machine Learning and Data Mining in Pattern Recognition. Springer, 2024, pp. 255–270. [Online]. Available: [https://link.springer.com/chapter/10.1007/978-3-031-80334-5\\_19](https://link.springer.com/chapter/10.1007/978-3-031-80334-5_19)
- [4] R. Xu, et al., "Interpretable Multimodal Large Language Models," arXiv preprint arXiv:2412.02104, 2024. [Online]. Available: <https://arxiv.org/abs/2412.02104>
- [5] T. Li, et al., "A Survey of Financial AI: Architectures, Advances, and Open Challenges," arXiv preprint arXiv:2411.12747, 2024. [Online]. Available: <https://arxiv.org/pdf/2411.12747>
- [6] J. Wu, et al., "A Survey of Large Language Models in Finance (FinLLMs)," arXiv preprint arXiv:2402.02315, 2024. [Online]. Available: <https://smallake.kr/wp-content/uploads/2024/10/2402.02315v1.pdf>
- [7] X. Wang, et al., "Financial Large Language Models: Techniques and Benchmarks," arXiv preprint arXiv:2411.06284, 2024. [Online]. Available: <https://arxiv.org/pdf/2411.06284>
- [8] X. Wang, et al., "Financial Large Language Models: Techniques and Benchmarks," arXiv preprint arXiv:2411.06284, 2024. [Online]. Available: <https://arxiv.org/pdf/2411.06284>
- [9] X. Dong, et al., "Large Language Models in Finance and Beyond," MIT Media Lab Research Paper, 2024. [Online]. Available: <https://web.media.mit.edu/~xdong/paper/jpm24b.pdf>
- [10] L. Zhang and H. Chen, "Multimodal AI: Challenges and Opportunities," Cognitive Computation and Systems, vol. 6, no. 2, pp. 112–125, 2024. [Online]. Available: <https://link.springer.com/article/10.1007/s44267-024-00065-8>



- [11] M. Shoman, et al., “Explainable Multimodal AI Systems,” AI, vol. 3, no. 3, pp. 1–20, 2024. [Online]. Available: <https://www.mdpi.com/2674-1032/3/3/25>
- [12] Y. Liu, et al., “Trustworthy AI in Finance: Foundations and Challenges,” in Proc. IEEE Conf. Artificial Intelligence, 2024, pp. 1–9. [Online]. Available: <https://ieeexplore.ieee.org/abstract/document/10521640/>
- [13] S. Kumar and P. Rao, “Foundation Models in Finance: Trends and Applications,” SSRN Electronic Journal, 2024. [Online]. Available: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=5196206](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5196206)
- [14] A. Sharma, et al., “AI-driven Systems for Finance and Governance,” Int. J. Web Grid Services, vol. 20, no. 1, pp. 22–38, 2024. [Online]. Available: <https://www.inderscienceonline.com/doi/abs/10.1504/IJWGS.2024.143177>
- [15] P. Martin, “Financial NLP with Large Language Models,” HAL Open Science, 2025. [Online]. Available: <https://hal.science/hal-04999764/>
- [16] Y. Wang, et al., “Foundation Models for Finance,” arXiv preprint arXiv:2408.01319, 2024. [Online]. Available: <https://arxiv.org/abs/2408.01319>
- [17] Z. Li, et al., “Large Language Models for Biomedicine and Finance,” The Innovation, vol. 6, no. 2, pp. 122–135, 2025. [Online]. Available: [https://www.cell.com/the-innovation/fulltext/S2666-6758\(25\)00151-1](https://www.cell.com/the-innovation/fulltext/S2666-6758(25)00151-1)
- [18] J. Park, et al., “Recent Advances in Financial AI,” arXiv preprint arXiv:2506.12958, 2025. [Online]. Available: <https://arxiv.org/pdf/2506.12958>
- [19] A. Singh, et al., “AI for Financial Decision Support,” in Proc. IEEE Int. Conf. Intelligent Systems, 2025, pp. 1–8. [Online]. Available: <https://ieeexplore.ieee.org/abstract/document/10829820>
- [20] S. Tan, et al., “Financial Document Analysis with Large Language Models,” arXiv preprint arXiv:2401.03428, 2024. [Online]. Available: <https://arxiv.org/abs/2401.03428>
- [21] F. Hu, et al., “LLMs in Healthcare and Finance,” Molecular Therapy – Nucleic Acids, vol. 34, pp. 212–225, 2024. [Online]. Available: [https://www.cell.com/molecular-therapy-family/nucleic-acids/fulltext/S2162-2531\(24\)00142-2](https://www.cell.com/molecular-therapy-family/nucleic-acids/fulltext/S2162-2531(24)00142-2)
- [22] M. Shoman, “Large Language Models: Applications, Challenges, Limitations and Future Prospects,” ResearchGate Preprint, 2024. [Online]. Available: <https://www.researchgate.net/profile/Maged-Shoman-2/publication/383818024>
- [23] R. Gayathri, et al., “Unpublished Manuscript on AI in Finance,” Personal Research Document, 2025.
- [24] D. Zhao, et al., “Financial Multimodal AI,” arXiv preprint arXiv:2503.21422, 2025. [Online]. Available: <https://arxiv.org/abs/2503.21422>
- [25] K. Nair and A. Thomas, “AI for Financial Networks,” Computer Communications, 2025. [Online]. Available: <https://www.sciencedirect.com/science/article/abs/pii/S1389128625003469>
- [26] Y. Zhou, et al., “Benchmarking Foundation Models in Finance,” arXiv preprint arXiv:2402.01801, 2024. [Online]. Available: <https://arxiv.org/abs/2402.01801>
- [27] C. Martin, et al., “Financial AI in Practice,” HAL Open Science, 2025. [Online]. Available: <https://hal.science/hal-05009116/>
- [28] S. Lopez, “Cross-Modal Reasoning in Finance,” HAL Open Science, 2025. [Online]. Available: <https://hal.science/hal-05020829/>
- [29] Georgia Southern University, “Large Language Models in Finance,” Electronic Theses and Dissertations, 2025. [Online]. Available: <https://digitalcommons.georgiasouthern.edu/etd/2715/>





INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  [ijirset@gmail.com](mailto:ijirset@gmail.com)



[www.ijirset.com](http://www.ijirset.com)

Scan to save the contact details