**Literature Review: The Impact of Large Language Models (LLMs) in Finance**

The integration of Artificial Intelligence (AI) and Machine Learning (ML) in finance has brought about a transformative shift in how financial data is processed and analyzed. At the forefront of this revolution are Large Language Models (LLMs), which have significantly expanded the role of AI in finance, particularly in Natural Language Processing (NLP) (Wei et al., 2022; Srivastava et al., 2022; Touvron et al., 2023). LLMs represent a substantial breakthrough in understanding and generating human-like text by utilizing deep learning techniques, particularly transformer architectures (Vaswani et al., 2017; Dong et al., 2023). These models are capable of processing vast amounts of data, including unstructured text, capturing semantic relationships between words and phrases (Adnan & Akbar, 2019). Additionally, LLMs can handle visual as well as multimodal data, learning semantic relationships across different modalities (Awais et al., 2023; Zhao et al., 2023). This advancement has profoundly impacted finance, where vast volumes of unstructured textual data—such as financial reports, news articles, and social media posts—are generated daily.

LLMs are being deployed for a wide range of critical financial tasks, including risk modeling, portfolio management, algorithmic trading, fraud detection, loan and insurance underwriting, customer service, and asset price prediction (Sen, Mehtab, & Engelbrecht, 2021). While some of these applications have been successfully implemented, others face business and operational challenges, particularly related to data privacy, model interpretability, and regulatory compliance. LLMs such as BloombergGPT, a 50-billion-parameter model trained on a diverse financial corpus, have transformed financial NLP tasks like news classification, entity recognition, and question answering (Wu et al., 2023). By leveraging vast financial datasets, BloombergGPT significantly enhances customer service by efficiently handling queries and providing high-quality financial advisory. Given the sensitive nature of financial information, it is imperative to employ robust measures like data encryption, data redaction, and stringent data protection policies to ensure compliance with data privacy regulations (Hadi et al., 2024).

This review aims to explore the current knowledge on LLMs in finance, evaluate the existing literature, and identify both the strengths and limitations of these models in real-world applications. It seeks to address key trends, challenges, and potential gaps in the use of LLMs, providing an overview of how these models are reshaping financial practices. The audience for this review includes researchers, financial practitioners, and AI developers interested in understanding the current and potential future impact of LLMs on the financial industry.

By focusing on key applications—such as sentiment analysis, fraud detection, customer service, and risk management—this review highlights both the potential and challenges of LLMs in transforming the financial sector. It also emphasizes the ethical and operational hurdles that must be addressed, including the risks of hallucination in model outputs, the handling of sensitive financial data, and the need for greater transparency and regulatory compliance. In doing so, this review provides a critical foundation for understanding how LLMs are likely to influence the future of finance.

**Applications of Large Language Models in Finance**

**Financial Market Prediction and Sentiment Analysis**

One of the most widely researched and applied uses of LLMs in finance is financial market prediction through sentiment analysis. Financial sentiment analysis entails extracting and interpreting opinions and emotions from vast amounts of unstructured data, such as news articles, social media posts, and financial reports. LLMs like FinBERT, GPT-3, and BloombergGPT have demonstrated remarkable success in this area, enabling more accurate stock price predictions and improved market forecasting.

Kirtac and Germano (2024) highlight the predictive power of FinBERT and GPT-3-based models in analyzing financial sentiment from sources such as corporate earnings reports and market news. By leveraging advanced natural language understanding, these models can provide real-time insights into market sentiment, which is crucial for financial decision-making. For example, by analyzing how news of geopolitical events or corporate performance influences market sentiment, LLMs can predict stock movements more accurately than traditional statistical models.

Moreover, LLMs’ capacity to process vast amounts of textual data in real time provides significant advantages in high-frequency trading environments, where timely insights can lead to profitable trading decisions. BloombergGPT, trained specifically on a financial corpus, enhances this capability by capturing subtle shifts in sentiment across multiple financial sources, including press releases, market updates, and media coverage (Kang & Liu, 2023). Financial institutions use these insights to inform algorithmic trading strategies, which have proven more responsive to real-time sentiment changes than human-driven analyses.

LLMs also contribute to the creation of sentiment-driven trading strategies, where algorithms execute trades based on shifts in market sentiment. This has become increasingly important in volatile markets where rapid fluctuations in investor sentiment can drive significant price movements. As these models become more sophisticated, financial institutions are adopting them to refine risk management strategies, improve portfolio diversification, and capture alpha in both traditional and alternative investment markets.

**Fraud Detection and Anti-Money Laundering (AML)**

LLMs have been transformative in fraud detection and anti-money laundering (AML) by enabling the identification of anomalous patterns in transactional data that signal fraudulent behavior. Financial institutions handle massive volumes of transactions daily, and LLMs’ ability to understand patterns within this data and detect irregularities far surpasses traditional rule-based systems. Fraud detection tasks that previously required human analysts to manually sift through data are now automated using LLMs, allowing for real-time, proactive identification of fraud.

LLMs, with their deep learning architectures, can extract meaning from transaction descriptions, behavioral patterns, and even communication records between individuals involved in fraudulent activities. Ahmed et al. (2022) note that LLMs outperform conventional systems by identifying fraud patterns that are subtle and contextually complex. For instance, LLMs can detect money laundering schemes by cross-referencing customer transaction histories with unstructured external data, such as emails or legal documents, to identify suspicious behavior.

The application of LLMs in AML goes beyond fraud detection to include regulatory compliance automation. By understanding and analyzing legal texts, financial disclosures, and regulatory guidelines, LLMs can automate the process of ensuring compliance with international standards. Lee et al. (2024) report that LLMs, particularly those adapted for financial applications, have greatly reduced the time and cost associated with AML efforts while enhancing the accuracy of compliance reporting. This makes them indispensable tools for global financial institutions that must navigate complex regulatory environments.

**Customer Service and Financial Advisory**

In customer service and financial advisory, LLMs have brought about a paradigm shift by enabling more intelligent, human-like interactions between clients and financial institutions. Financial institutions increasingly deploy LLM-powered chatbots and virtual assistants to handle routine customer inquiries, offer financial advice, and assist users through financial decisions. LLMs like ChatGPT and BloombergGPT allow these systems to understand the context of client queries, making interactions more efficient and personalized.

These models are capable of offering tailored financial advice based on customer history, preferences, and goals. Khan and Umer (2024) describe how ChatGPT has been instrumental in improving customer satisfaction by providing timely, accurate, and personalized responses to customer queries. LLMs’ ability to deliver high-quality customer service not only strengthens client relationships but also alleviates the workload of human customer service representatives, enabling them to focus on more complex or high-priority tasks.

In addition to answering routine queries, LLMs can assist with more sophisticated tasks, such as recommending investment products, guiding clients through portfolio diversification, or generating real-time financial insights. BloombergGPT, specifically trained on financial texts, is capable of providing more specialized advice by interpreting financial reports, market trends, and company performance data (Kang & Liu, 2023). This has proven especially valuable for high-net-worth clients and institutional investors, who require more detailed and expert-level guidance.

Moreover, LLMs have the potential to democratize financial advisory services by making them more accessible to retail investors who may not have the resources for personalized financial advice from human advisors. By providing high-quality, automated financial advice, LLMs enable individuals with smaller portfolios to access insights previously only available to more affluent investors.

**Risk Management and Portfolio Optimization**

LLMs are playing a critical role in risk management by providing advanced insights into market risk, credit risk, and operational risk. Financial institutions rely on LLMs to assess risk exposure by analyzing diverse datasets, including financial statements, economic forecasts, and market performance indicators. These models excel at detecting emerging risks by identifying patterns and correlations .

According to Goldberg (2024), LLMs like GPT-4 have proven especially valuable in automating the generation of risk reports that integrate data from multiple sources, allowing risk managers to act swiftly in volatile market conditions. For example, LLMs can identify early warning signals in company earnings reports that suggest potential defaults or liquidity crises. By automating the analysis of such reports, LLMs help institutions anticipate risks before they escalate, allowing for more proactive risk management strategies.

In portfolio optimization, LLMs are used to develop data-driven strategies that maximize returns while minimizing risks. These models analyze historical market data, investor sentiment, and real-time economic indicators to suggest optimal asset allocation. Lee et al. (2024) point out that LLMs enable portfolio managers to process complex, multifactorial data sets, allowing for more informed decisions on asset allocation and risk diversification. Consequently, LLMs strengthen institutional investors’ ability to navigate market fluctuations and seize emerging opportunities.

**Multimodal Data Processing and Numerical Reasoning**

LLMs are also increasingly used to process multimodal data, combining textual, numerical, and visual data (e.g., tables and graphs) to provide comprehensive financial insights. Models such as DOCMATH-EVAL have been developed specifically to assess the numerical reasoning capabilities of LLMs when interpreting financial documents (Zhao et al., 2023). This is particularly important in finance, where accurate numerical interpretation is essential for tasks such as tax reporting, auditing, and regulatory compliance.

While LLMs have demonstrated significant proficiency in processing textual data, their ability to handle complex numerical reasoning tasks still poses challenges. For example, models like GPT-4 can handle basic arithmetic and table interpretation but often struggle with more advanced financial calculations or the interpretation of interdependent datasets over long periods. Despite these challenges, the integration of LLMs with other AI tools, such as deep learning models for numerical analysis, offers promising advancements for the future of multimodal data processing in finance.

**Critical Evaluation of the Literature**

The existing literature offers substantial evidence of LLMs’ transformative potential in finance. Their ability to process vast amounts of data and provide real-time insights across multiple domains—market prediction, fraud detection, customer service, and risk management—is clear. However, several limitations remain. Hallucination, the generation of incorrect or fabricated information, poses a major risk when using LLMs in critical financial decision-making. Additionally, the black-box nature of LLMs complicates their interpretability, making it difficult for financial institutions to meet regulatory requirements for transparency and accountability.

Moreover, while LLMs excel in processing textual data, their performance in handling complex numerical reasoning tasks still lags behind that of human experts. The ethical concerns associated with biased or inaccurate recommendations further complicate their widespread adoption. Despite these challenges, LLMs have made significant strides in transforming financial services, and future developments are likely to mitigate many of these limitations.

**Future Directions**

Future research should focus on improving the explainability of LLMs to ensure compliance with regulatory standards and increase trust in AI-driven financial systems. Developing domain-specific models like BloombergGPT and FinBERT that are tailored to the needs of the finance industry will also be critical. Addressing ethical concerns such as bias and misinformation through better training protocols and hybrid AI-human decision-making systems will ensure that LLMs are used responsibly in finance.

**References**

Adnan, K., & Akbar, R. (2019). An analytical study of information extraction from unstructured and multidimensional big data. *Journal of Big Data, 6*(1), 1–38.

Awais, M., Naseer, M., Khan, S., Anwer, R. M., Cholakkal, H., Shah, M., Yang, M.-H., & Khan, F. S. (2023). Foundational models defining a new era in vision: A survey and outlook. *arXiv preprint arXiv:2307.13721*.

Ahmed, S., Alshater, M. M., El Ammari, A., & Hammami, H. (2022). Artificial intelligence and machine learning in finance: A bibliometric review. *Research in International Business and Finance, 61*, 101646.

Dong, Z., Tang, T., Li, L., & Zhao, W. X. (2023). A survey on long text modeling with transformers. *arXiv preprint arXiv:2302.14502*.

Goldberg, A. (2024). AI in Finance: Leveraging Large Language Models for Enhanced Decision-Making and Risk Management. *Social Science Journal for Advanced Research, 4*(4), 33–40. https://doi.org/10.5281/zenodo.13299843

Kang, H., & Liu, X. Y. (2023). Deficiency of Large Language Models in Finance: An Empirical Examination of Hallucination. *Columbia University and Rensselaer Polytechnic Institute*.

Kirtac, K., & Germano, G. (2024). Sentiment trading with large language models. *Finance Research Letters, 62*, 105227.

Khan, M. S., & Umer, H. (2024). ChatGPT in finance: Applications, challenges, and solutions. *Heliyon, 10*, e24890.

Lee, J., Stevens, N., Han, S. C., & Song, M. (2024). A Survey of Large Language Models in Finance (FinLLMs). *FinLLMs*.

Sen, J., Mehtab, S., & Engelbrecht, A. (2021). *Machine Learning: Algorithms, Models and Applications*. Artificial Intelligence, Volume 7. BoD – Books on Demand.

Srivastava, A., Rastogi, A., Rao, A., Shoeb, A. A. M., Abid, A., Fisch, A., Brown, A. R., Santoro, A., Gupta, A., & Garriga-Alonso, A. (2022). Beyond the imitation game: Quantifying and extrapolating the capabilities of language models. *arXiv preprint arXiv:2206.04615*.

Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M.-A., Lacroix, T., Rozière, B., Goyal, N., Hambro, E., & Azhar, F. (2023). Llama: Open and efficient foundation language models. *arXiv preprint arXiv:2302.13971*.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. In *Advances in Neural Information Processing Systems* (pp. 5998–6008).

Wei, J., Tay, Y., Bommasani, R., Raffel, C., Zoph, B., Borgeaud, S., Yogatama, D., Bosma, M., Zhou, D., & Metzler, D. (2022). Emergent abilities of large language models. *arXiv preprint arXiv:2206.07682*.

Wu, S., Irsoy, O., Lu, S., Dabravolski, V., Dredze, M., Gehrmann, S., Kambadur, P., Rosenberg, D., & Mann, G. (2023). BloombergGPT: A large language model for finance. *arXiv preprint arXiv:2303.17564*.

Zhao, Y., Lin, Z., Zhou, D., Huang, Z., Feng, J., & Kang, B. (2023). BuboGPT: Enabling visual grounding in multi-modal LLMs. *arXiv preprint arXiv:2307.08581*.

Zhao, Y., Long, Y., Liu, H., Kamoi, R., Nan, L., Chen, L., Liu, Y., Tang, X., Zhang, R., & Cohan, A. (2024). DOCMATH-EVAL: Evaluating Math Reasoning Capabilities of LLMs in Understanding Financial Documents. *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 16103–16120.