Publisher

www.singhpublication.com

E-ISSN:2583-2468

Applied Science and Engineering Journal for Advanced Research

2025 Volume 4 Number 1 January

Research Article

Fraud Detection

LLM for Financial Services: Risk Analysis and Fraud Detection

Sudhakar VM[1*]

DOI:10.5281/zenodo.14928807

[1*] Vinoth Manamala Sudhakar, Sr. Data Scientist (Independent Researcher), Cloud Software Group Inc., Austin, Texas, United States of America.

The financial service industry is increasingly suspected by risk management and complicated frauds, because of traditional methods, such as rules based on rules, becomes become Not enough to combat evolutionary threats. This study discovers the potential of large language models (LLM), including GPT-3 and Finbert, to improve risk analysis and fraud detection in the financial sector. LLM, capable of processing structured and non -structured data, provides improvement in detecting models and abnormalities between trading newspapers, customer interaction and talent reports main. A quantitative comparative comparative research design, financial data analysis can access the public and compare LLM performance with traditional systems. Main performance measures - Prediction Accuracy, False Positive Rate, Processing Time, and Fraud Detection Rate- are used to evaluate the effectiveness of the models. The results show the significant potential of LLM to improve financial risk management and detect fraud, provide an effective, accurate and developed approach to modern financial institutions.

Keywords: large language models (llms), gpt-3, finbert, risk analysis, fraud detection, financial services

Corresponding Author	How to Cite this Article	To Browse
Vinoth Manamala Sudhakar, Sr. Data Scientist (Independent Researcher), Cloud Software Group Inc., Austin, Texas, United States of America. Email: vinoth.manamala@cloud.com	Sudhakar VM, LLM for Financial Services: Risk Analysis and Fraud Detection. Appl. Sci. Eng. J. Adv. Res.. 2025;4(1):65-70. Available From https://asejar.singhpublication.com/index.php/ojs/article/view/132	

Manuscript Received	Review Round 1	Review Round 2	Review Round 3	Accepted
2024-12-26	2025-01-11			2025-01-27
Conflict of Interest	**Funding**	**Ethical Approval**	**Plagiarism X-checker**	**Note**
None	Nil	Yes	3.77	

1. Introduction

The financial service industry is an important pillar of the global economy, but it faces a series of challenges in risk management and fraud. Financial fraud programs become more sophisticated and develop rapidly, traditional methods such as rules based on rules and manual control are increasingly incomplete. These methods often have difficulty detecting new frauds or accurately assessing risks for their dependence on static and predetermined rules. As the complexity of financial transactions increases and the amount of data increases, there is a clear demand for more advanced and dynamic tools that can follow these challenges.

large language models (LLM), a collection of artificial intelligence (AI), has become a promising solution for these urgent issues in the financial sector. LLM, formed on large data sets and has the ability to understand and create human documents, can process structured and structured data, commonly in financial services. By analyzing trading newspapers, customer interactions and financial statements, LLM can identify models and abnormalities that may indicate risks or fraudulent activities. The ability to manage and explain this complex data allows financial institutions to improve their decision -making processes, improve fraud detection systems and minimize risks in real time.

1.1 Applications of LLMs in Risk Analysis and Fraud Detection:

- **Risk Analysis:** LLM can automate the analysis of large data sets, making more accurate and timely assessments on credit and market risks. By combining data from articles, reports and social media, these models can detect early warning signs of potential system risks, making it one. Pricing tools for financial institutions.
- **Fraud Detection:** LLM improves the detection of fraud by identifying complex models in trading data showing fraudulent behavior. The ability to process data is structured and without their structure, such as email or customer messages on social networks, giving them an advantage in discovering frauds that cannot be played without being played. Currently by traditional systems.
- **Comply with the Regulations:** LLM can automate the inspection of financial documents and ensure compliance with the continuous development regulations. This capacity is especially useful for financial institutions operating in many legal areas, in which the executives are significantly different.

2. Literature Review

Fan, M. (2024) discovered applications, challenges and strategic methods for large language models (LLM) in the banking sector. The first article built an analysis framework that combined the commercial value chain with technical sizes, identifying and evaluating LLM's extended application scenarios in the bank. This document was then transmitted to the main challenges, such as data security and security, modeling ability, algorithm deviations, and system integration, and provided corresponding strategies. In addition, it provided strategic recommendations for banks to implement LLM, emphasizing system planning, building data ecosystems, and compliance management. This article aimed at providing both theoretical information and practical advice to banks that sought to take advantage of LLM for technology conversion and competitive advantage.

Luca, C. (2023) discovered the optimization of large language models (LLM) to assess financial risks in credit cooperatives, focusing on the ability to process important and non-structured financial data. Predicted the hidden risk. The process of assessing financial risks in credit cooperatives played an essential role in identifying potential dangers and ensuring their long-term stability. However, traditional risk assessment models were often not flexible and accurately needed to manage the complexity of modern financial systems. We proposed a method in which LLM, especially well-treated models such as Bert and GPT, were applied to predict default risks, frauds, and credit symbols. By taking advantage of historical financial data, models were trained to improve the accuracy of predictions related to traditional statistical methods. The results showed that LLM significantly improved the accuracy and effectiveness of risk assessments in credit cooperatives, providing a promising alternative to inherited models. This article discussed the meaning of LLM integration in their credit unions and its potential to revolutionize the context of financial risk assessment.

Lakkaraju, et al. (2023)decided to study how those systems operated in the field of personal finance, which included finance as an overall goal of banks for decades.

We checked the LLM, Chatgpt, and Bard chatbot, which were widely used, and compared their performance with Safe finance, a chatbot based on rules built on the RASA platform. The comparison was between two important tasks: discovering products and multi-product interaction, in which the product referred to bank products such as credit cards, deposit certificates, and account checking. With this study, we provided interesting information about the effectiveness of chatbots in financial advice and their ability to provide fair treatment among different user groups. We noted that Bard and TATGPT might have made mistakes in the recovery of basic online information, the answers they created were not suitable between different user groups and they could not be called to have gotten reasons related to banking products. On the other hand, although their ability to generalize limitations, chatbots based on rules such as Safe finance provided safety and reliable feedback for users that could be searched at their original source. Overall, although the results of chatbots based on LLM were popular and reasonable, there were still important shortcomings in providing coherent and reliable financial information.

Park, T. (2023) presented a multi-agent framework based on a large language model (LLM) that was designed to improve the detection of abnormalities in the financial market data, offering long-term challenges to manually verify the heterosexuals. Usually was created by the system. The framework of operating a cooperation network of AI dealers, each worked specialized in separate functions, especially data conversion, expert analysis through web research, using physical knowledge institutional or the role of consolidation and reporting of reports. By coordinating those agents aimed at a common goal, the frame provided a complete and automatic approach to confirm and explain the financial data anomalies. I analyzed the S&P 500 index to prove the capacity of the framework in improving efficiency, accuracy, and reducing human intervention in monitoring the financial market. The integration of autonomous AI functions with established analytical methods not only highlighted the effectiveness of the frame in detecting abnormalities, but also showed its wider applicability to it to be. supported financial market monitoring.

3. Research Methodology

3.1 Research Design

This study applies a quantitative comparative design, financial data analysis that can be accessed to the public through LLM such as GPT-3 and Finbert. Research comparing LLM with traditional systems based on rules using performance measures. A statistical analysis will be applied to evaluate the difference and evaluate the effectiveness of LLM in improving financial risk management and fraud detection.

3.2 Data Collection

For this study, data sets and financial cases can be accessed to the public used to prove LLM's applicability in financial services. The collected data includes trading newspapers, customer behavior, financial statements and past frauds detection files. The main sources include:

- **Transaction Logs:** They include details of financial transactions such as the amount, related accounts, horatages and market information.
- **Customer Interaction Data:** Records, including email, cat and comment, analysis for models that may indicate fraud or risk factors.
- **Financial Statements:** They include assessments, results status and audit records, providing context to assess the financial health of individuals and organizations.

This implies data cleaning, managing the lack of values and converting it into formats that adapt to LLMS analysis. The data used is not only history, but also includes progressive transactions to assess the actual time capabilities of LLM.

3.3 Model Selection

Research focuses on the following LLM, each LLM is chosen because of the ability to handle important and without structures and success in the tasks of natural language processing, especially in the financial sector:

- **GPT (Generative Pretrained Transformer):** It is known for the language of generation capacity and the understanding of text, GPT is used for evil analysis Non -structured text as customer comments, descriptions of transactions and market relationships.

- **Finbert:** A dedicated variant of Bert adapted to the financial sector, Finbert is used to be able to process financial data, including sensory analysis of financial and new relationships. , to help rank credit and detect fraud.

These models are selected according to their identification models, the ability to analyze data with structured and non -structural and their ability to financial applications.

3.4 Evaluation Metrics

The effectiveness of large linguistic models (LLM) in financial risks and fraud is evaluated by the main measures: Prediction Accuracy, False Positive Rate, Processing Time, and Fraud Detection Rate. These indicators measure the capabilities of models to determine risks and fraud with accuracy, reduce operating costs and provide real -time performance.

3.5 Data Analysis Techniques

The data analysis process includes some main steps to assess the effectiveness of large language models (LLM) in financial risk analysis and fraud detection. First, raw financial data and customer interactions are processed first, clean it and turn it into formats suitable for models. The LLM was then selected (GPT-3 and Finbert) formed on prepared data sets, with good adjustments to adapt to tasks such as risk analysis and fraud detection. After being trained, the models are tested on intangible data to evaluate their performance by using the set measures, such as prediction accuracy, fake positive rates and Fraud detection rate. Finally, the LLM results are compared to traditional systems based on rules to assess the improvement of accuracy, processing time and the ability to detect fraud.

4. Data Analysis

In this section, the effectiveness of conventional rule-based systems and Large Language Models (LLMs) such as GPT-3 and FinBERT in financial risk management and fraud detection is compared.

Table 1: Prediction Accuracy of LLM vs Traditional Methods

	Prediction Accuracy (%)	False Positive Rate (%)	Processing Time (s)
Traditional Rule-Based	78	15	4.5
LLM (GPT-3 / FinBERT)	92	5	2.0

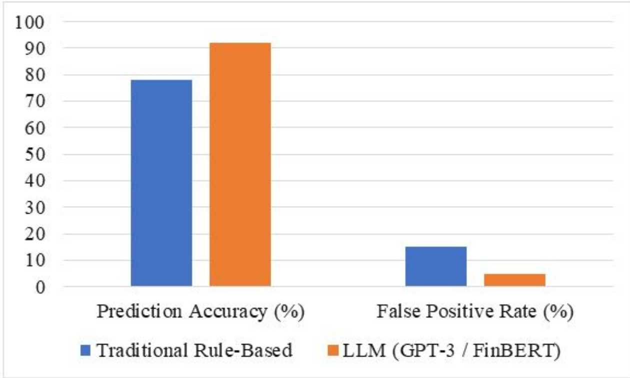

Figure 1: Prediction Accuracy of LLM vs Traditional Methods

The results presented in Table 1 highlighted the significant improvement of the accuracy of the prediction, the positive ratio and the treatment time when using important language models (LLM) such as GPT- 3 and Finbert compared to traditional methods based on rules in credit risk assessment. LLM achieved 92%prediction, this is a significant increase compared to the inhibition of 78 ° observed with traditional methods. This shows that LLM can provide more accurate and reliable risk assessments by analyzing a large amount of data and identifying models that may be overlooked by rules based on rules. In addition, the fake positive rate with LLM is significantly lower at 5%, compared to 15% for traditional methods, this indicates that LLM is more effective in the exact classification of records. Risks without reporting innocent transactions are high risks. In addition, LLM provides a more effective handling time of only 2.0 seconds, which significantly reduces the time needed to assess risks compared to 4.5 seconds for traditional methods. These results emphasize the potential of LLM to streamline activities, improve accuracy and provide faster and more reliable credit risk assessments.

Table 2: Comparison of Fraud Detection Performance

	Detection Rate (%)	False Positive Rate (%)	Processing Time (s)
Rule-Based System	82	20	6.5
LLM (GPT-3 / FinBERT)	95	3	3.2

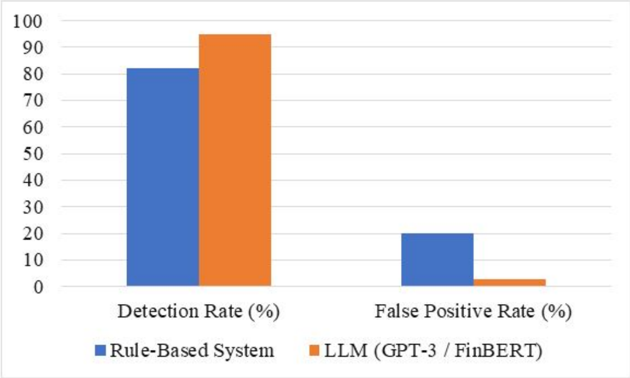

Figure 2: Comparison of Fraud Detection Performance

The results of Table 2 clearly prove that large language models (LLM) such as GPT-3 and Finbert surpassed traditional systems based on rules in detecting frauds for banking transactions. The detection rate for LLM is 95%, compared to 82% for rules based on rules, this indicates that LLM is much more effective in determining fraudulent activities. In addition, the fake positive rate with LLM is much lower, only at 3%, compared to 20% with rules based on rules. This shows that LLM is better to distinguish legal and fraudulent transactions, reducing opportunities to point out the wrong activities. In addition, LLM provides faster treatment time than 3.2 seconds, significantly improved effectiveness compared to 6.5 seconds at the request of rules based on rules. These results emphasize that the LLM not only improves the accuracy and reliability of fraud but also allows the reality of real and real decisions, making it an appropriate choice. more suitable for modern financial institutions.

Table 3: Performance Comparison of LLMs (GPT-3 / FinBERT) and Traditional Methods (Financial Risk Management)

	Prediction Accuracy (%)	False Positive Rate (%)	Fraud Detection Rate (%)	Processing Time (s)
Traditional Rule-Based	78	15	82	4.5
LLM (GPT-3 / FinBERT)	92	5	95	2.0

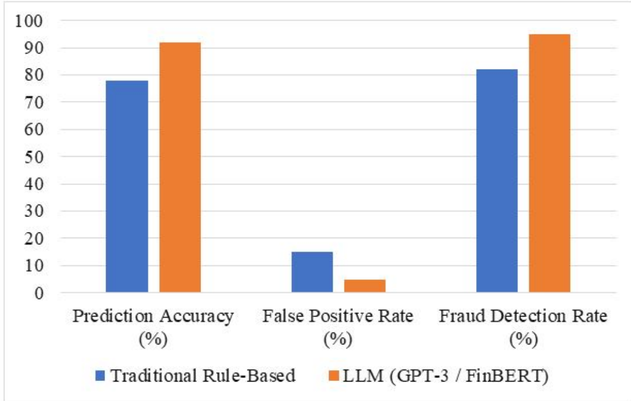

Figure 3: Performance Comparison of LLMs (GPT-3 / FinBERT) and Traditional Methods (Financial Risk Management)

The results presented in Table 3 show a clear advantage of large language models (LLM) like GPT-3 and Finbert based on traditional methods based on rules in talent risk management. main and discover fraud. LLM has achieved a 92% prediction, a significant improvement compared to the accuracy of 78% of the traditional method. This reflects the higher power of LLM in providing specific risk assessments and fraud. In addition, the fake positive rate with LLM is much lower at 5%, compared to 15% for traditional systems, minimizing unnecessary signals of legal transactions. Treatment time with LLM is also significantly faster after 2.0 seconds, bringing higher efficiency for real applications. Finally, the LLMS fraud detection rate is 95% higher, showing the ability to determine more effectively frauds compared to the detection rate of 82% of traditional rules based on rules. These results prove that the LLM provides accuracy, efficiency and reliability increased in financial risk management and frauds detection tasks.

5. Conclusion

This study emphasized the significant advantages of large language models (LLM) such as GPT-3 and Finbert in financial risk analysis and fraud detection compared to traditional rules based on rules. The results showed that LLM achieved a significant higher prediction and prediction rate detection rate, as well as significantly reducing the positive rate and faster processing time. These improvements are motivated by the ability of LLM to process structured and unconcerned data, allowing them to identify complex models and detect abnormalities in real time.

Given that financial institutions face the number of sophisticated frauds and increasing dynamic risks, the implementation of LLM provides a powerful solution to improve operational efficiency and reduce errors of People and make accurate decisions focused on data. In addition, the ability to adapt to the development of the financial landscape and the executive director stipulates their positioning as an essential tool to imply financial services. The study emphasized that the application of LLM -based technologies can cause significant improvements in risk management, fraud and compliance detection, which makes it an essential asset in effort. Continuous force to ensure financial systems in the world.

References

1. Desta. (2020). Analysis of the Ethiopian banks response. *Bahir Dar University School of Law LLM Programme in Business and Corporate Law: A Research for Seminar on Contemporary Business Law Issues*.

2. Bakumenko, K. Hlaváčková-Schindler, C. Plant, & N. C. Hubig. (2024). *Advancing anomaly detection: Non-semantic financial data encoding with LLMs*. arXiv preprint arXiv:2406.03614.

3. Cao, X. Yao, W. Gong, Y. Li, Y. Pan, J. Zhang, ... & C. Zhao. (2024). *LLMs for insurance: Opportunities, challenges and concerns*.

4. Chadalavada, T. Huang, & J. Staddon. (2024). *Distinguishing scams and fraud with ensemble learning*. arXiv preprint arXiv:2412.08680.

5. Zhao, Z. Liu, Z. Wu, Y. Li, T. Yang, P. Shu, ... & T. Liu. (2024). *Revolutionizing finance with LLMs: An overview of applications and insights*. arXiv preprint arXiv:2401.11641.

6. Chakraborty, W. Xia, A. Majumder, D. Ma, W. Chaabene, & N. Janvekar. (2024). Detoxbench: *Benchmarking large language models for multitask fraud & abuse detection*. arXiv preprint arXiv:2409.06072.

7. Xu, H. Wang, Y. Zhong, L. Qin, & Q. Cheng. (2024). Predict and optimize financial services risk using ai-driven technology. *Academic Journal of Science and Technology, 10*(1), 299-304.

8. Jimma. (2022). LLM in human rights and criminal law. *Doctoral Dissertation, Jimma University*.

9. Lakkaraju, S. E. Jones, S. K. R. Vuruma, V. Pallagani, B. C. Muppasani, & B. Srivastava. (2023). LLMs for financial advisement: A fairness and efficacy study in personal decision making. in *Proc. Fourth ACM International Conference on AI in Finance*, pp. 100-107.

10. Luca. (2024). *Optimizing large language models for financial risk assessment in credit unions*.

11. Lui, & N. Ryder, Eds. (2021). *FinTech, artificial intelligence and the law: Regulation and crime prevention*. Routledge.

12. Fan. (2024). LLMs in banking: Applications, challenges, and approaches. in *Proc. International Conference on Digital Economy, Blockchain and Artificial Intelligence*, pp. 314-321.

13. Pankajakshan, S. Biswal, Y. Govindarajulu, & G. Gressel. (2024). *Mapping LLM security landscapes: A comprehensive stakeholder risk assessment proposal*. arXiv preprint arXiv:2403.13309.

14. Park. (2024). *Enhancing anomaly detection in financial markets with an LLM-based multi-agent framework*. arXiv preprint arXiv:2403.19735.

15. W. Chang, S. Sarkar, S. Mitra, Q. Zhang, H. Salemi, H. Purohit, ... & C. T. Lu. (2024). Exposing LLM vulnerabilities: Adversarial scam detection and performance. in *Proc. 2024 IEEE International Conference on Big Data (BigData)*, pp. 3568-3571.