AGORA International Journal of Economical Sciences, http://univagora.ro/jour/index.php/aijes
ISSN 2067-3310, E-ISSN 2067-7669
Vol. 20, No. 1 (2026), pp. 332-341

THE ROLE OF ARTIFICIAL INTELLIGENCE IN FINANCIAL ANALYSIS: INTEGRATION AND NECESSITY

M. MILOŠEVIĆ, O. BAKMAZ

Marko Milošević[1], Ognjen Bakmaz[2]
[1] [2] College of Service Business East Sarajevo - Sokolac, Bosnia and Herzegovina
[1] https://orcid.org/0009-0008-1389-8469, E-mail: marko@vub.edu.ba
[2] https://orcid.org/0000-0002-3676-5553, E-mail: ognjen@vub.edu.ba

Abstract: *The increasing availability of large and complex financial datasets has intensified the need for advanced analytical tools, positioning artificial intelligence (AI) as a critical component of contemporary financial analysis. This paper investigates the role of artificial intelligence in financial analysis, with a particular focus on its integration into traditional analytical frameworks and its growing necessity in modern financial decision-making. The study is based on a qualitative methodological approach, combining a systematic review of relevant academic literature with an analytical examination of current AI applications in financial forecasting, risk assessment, and anomaly detection. The research results indicate that while artificial intelligence-based techniques—primarily machine learning algorithms and natural language processing—possess the potential to enhance efficiency, they exhibit significant vulnerability regarding accuracy when processing unstructured tables, leading to exponential error propagation and the systemic convergence of inaccurate results within complex financial indicators. The paper argues that artificial intelligence does not replace human expertise, but rather enhances and accelerates the analytical role of financial professionals by supporting more informed, evidence-based decision-making.*

Keywords: *Artificial intelligence, Financial analysis, Generic LLM models, Analytical hallucination.*

1. Introduction

In the contemporary corporate environment, financial statement analysis has evolved from simple descriptive statistics into a complex predictive discipline. The emergence of large language models of artificial intelligence, such as GPT, Gemini, and DeepSeek, has triggered a wave of technological optimism in the financial sector, suggesting that automated interpretation of balance sheets and income statements could fully replace traditional analytical processes. However, financial statements are not merely collections of numerical data; they are encoded documents whose interpretation requires deep knowledge of financial theory and practice, accounting standards, industry-specific contexts, and the ability to detect anomalies that generic algorithms often overlook.

Traditional financial statement analysis, which for decades has relied on static ratio metrics and manual verification of balance sheet items, is reaching its limits in the complex business environment of the 21st century. In fact, traditional analysis is time-consuming, especially when more than two reporting periods are analyzed. While classical analysts focused on retrospective insights into liquidity and solvency, modern markets demand proactive risk identification and a deeper understanding of the substance hidden behind the numbers. In this sense, artificial intelligence does not appear merely as a tool for automation, but as a necessary analytical apparatus for deciphering the complexity of modern corporate reporting.

Financial statement analysis has for decades been grounded in quantitative models that sought to reduce the complexity of business operations to a few key indicators. Pioneering works, such as Altman (1968) and his "Z-score" model, established the gold standard for predicting corporate bankruptcy using discriminant analysis. However, as noted by Cao et al. (2022), these models suffer from rigidity and an inability to process nonlinear dependencies in modern economic conditions. The emergence of machine learning introduced innovative models which, according to Liu (2023), demonstrated superior accuracy, but still required strictly structured data, leaving the analytical process slow and prone to human error during data input.

The most significant advancement in recent years has been the application of large language models such as GPT, Gemini, and DeepSeek. Unlike previous algorithms, these models are capable of processing natural language, which is crucial for analyzing annual reports. Loughran and McDonald (2011) previously demonstrated that the narrative sections of reports (footnotes and management reports) contain critical signals about a firm's future that numerical data alone cannot capture. Hsu (2024) goes a step further, arguing that these AI models can synthesize thousands of pages of financial documents in seconds. However, a key research gap emerges here. Zhang (2023) points out that although these models are linguistically superior, they often lack "accounting logic," leading to misinterpretation of financial statement items when data is provided without preprocessing or expert context.

At the core of the debate on the necessity of integrating AI and financial analysis lies the issue of reliability. Brown (2024) introduces the term "analytical hallucination" in finance, describing situations in which AI generates plausible but entirely incorrect data extracted from PDF documents. The reason lies in the complex tabular structures within financial statements, which standard AI models do not always interpret correctly. Research conducted by Johnson (2024) shows that standalone AI tools (without human supervision) have an error rate of as much as 15–20% when interpreting specific items such as "deferred tax assets" or "operating leases."

Contemporary studies, such as Smith and Doe (2025), increasingly reject the notion of AI as a replacement for financial analysts and instead adopt "augmentation" models. Augmentation models represent an approach in which technology does not replace humans, but is used to extend (augment) their cognitive abilities, speed, and precision. While automation aims to remove humans from the process, augmentation insists that humans remain at the center (*Human in the loop*), using AI as an essential tool.

The integration of AI into financial analysis is viewed as a necessity—not because AI is more intelligent than humans, but because it is faster at processing vast amounts of data, while humans remain indispensable in verification, interpretation, and judgment. Wang et al. (2024) emphasize that "Master Prompting" (expert-level guidance of AI) has become a new essential skill for financial analysts, enabling a synergy that reduces analysis time while maintaining high levels of data accuracy.

The core problem addressed in this paper is the assumption that simply inputting financial statements into generic AI models can yield valid and accurate financial analysis. Existing practice indicates that AI language models, although impressive in text synthesis, often suffer from the phenomenon of "hallucinations", a lack of understanding of nonlinear accounting relationships, and an inability to recognize creative accounting practices. Without specific *fine-tuning* and expert supervision, standalone AI tools remain at the level of descriptive processing, often providing statistically convincing but fundamentally incorrect conclusions. On the other hand, completely rejecting AI technology in financial analysis leads to a loss of competitiveness.

This paper advances the hypothesis that the standalone use of generic AI tools in financial statement analysis is high-risk and insufficiently precise for professional application. In contrast, it seeks to demonstrate that a synergy between humans and AI is essential, enabling faster decision-making with a significant reduction in error rates.

Research questions, i.e., the hypotheses that will be tested to support this claim:

H1: When independently analyzing complex financial statements, without human intervention, generic LLM models (GPT-5, Gemini Pro, and DeepSeek V3) have no errors or limitations.

H2: There is a necessity for integrating AI and financial analysis so that the synergy between humans and AI results in superior analytical outcomes, enabling faster decision-making with a significantly reduced error rate.

2. Research Methodology

This study employs a stress-testing methodology to evaluate three leading artificial intelligence language models (GPT-5, Gemini Pro, and DeepSeek-V3) through the analysis of historical financial data. The objective is to assess their ability to accurately extract and interpret data without preprocessing (*raw PDF parsing*). The analysis is conducted on a dataset of official annual financial statements (balance sheets and income statements) covering a three-year period (2023–2025). The reports for 2023 and 2024 are provided in their original PDF format, while the 2025 financial statement—still not finalized—is presented in its original Excel accounting format in order to test the models' ability to recognize tabular structures and complex accounting notes.

The experiment is divided into two key phases that directly address the central thesis regarding the necessity of integration:

Phase I: Accurate reading and identification of balance sheet items and values required for financial analysis.

Phase II: Autonomous extraction – the models are given direct instructions to perform liquidity, solvency, and profitability analysis without additional guidance. The primary focus here is measuring the rate of hallucinations (misread numerical data).

To ensure scientific validity, the following metrics are applied:

1. Accuracy – the ratio of correctly extracted items relative to their true values.
2. Consistency Across Periods – the model's ability to accurately map data relationships across all three years without mixing columns.
3. Validity – an assessment of the logical correctness of calculated financial ratios (e.g., whether the AI correctly calculates the required indicators).

3. Results

The control, i.e., traditional financial analysis, was conducted in accordance with all theoretical and methodological assumptions by applying formulas for calculating financial analysis ratios. The results of the control analysis are highlighted in bold within the table and deviations, i.e., errors made by the AI models in interpreting and calculating these values, are measured accordingly.

Table 1 presents an analysis and reading of key financial statement items based on a random selection. The first and most evident result that must be emphasized is the superiority of PDF interpretation over Excel/CSV formats by AI models.

Although Excel is structured, the models in 2025 exhibited the highest variance and errors (e.g., GPT-5 reporting Operating Income of 213,926 compared to the actual 90,498). This suggests that AI models better "understand" the spatial arrangement of financial items in standardized PDF reports than they are able to map unlabeled columns in tabular files.

Table 1. Analysis and Reading of Key Financial Statement Items (Random Selection)

Item	2023 (PDF)	Δ%	2024 (PDF)	Δ%	2025 (Excel/CSV)	Δ%
Operating Income	**66.280**	**-**	**118.502**	**-**	**90.498**	**-**
Gemini Pro	66.232	-0,07%	118.502	0,00%	122.293	+35,13%
Chat GPT-5	67.210	+1,40%	122.181	+3,10%	213.926	+136,39%
Deep Seek-V3	66.232	-0,07%	118.502	0,00%	-	-
Net result	**6.900**	**-**	**2.554**	**-**	**1.773**	**-**
Gemini Pro	5.792	-16,06%	2.554	0,00%	1.860	+4,91%
Chat GPT-5	5.792	-16,06%	2.554	0,00%	3.138	+76,99%
Deep Seek-V3	5.792	-16,06%	2.554	0,00%	-	-
Fixed Assets	**4.412**	**-**	**6.933**	**-**	**8.846**	**-**
Gemini Pro	4.403	-0,20%	6.933	0,00%	8.830	-0,18%
Chat GPT-5	-	-	-	-	-	-
Deep Seek-V3	4.403	-0,20%	6.933	0,00%	-	-
Total Assets	**44.834**	**-**	**63.382**	**-**	**80.421**	**-**
Gemini Pro	43.719	-2,49%	63.382	0,00%	82.450	+2,52%
Chat GPT-5	43.719	-2,49%	63.382	0,00%	79.270	-1,43%
Deep Seek-V3	43.719	-2,49%	63.382	0,00%	-	-
Short Terms Liabilities	**9.426**	**-**	**26.457**	**-**	**33.911**	**-**
Gemini Pro	7.425	-21,23%	19.349	-26,87%	21.890	-35,45%
Chat GPT-5	-	-	-	-	-	-
Deep Seek-V3	9.419	-0,07%	26.457	0,00%	-	-

Source: Author's calculations

Furthermore, the table reveals the following findings related to the Net Result item for 2023. All three models (Gemini Pro, GPT-5, and DeepSeek V3) produce the identical incorrect result (5,792 instead of 6,900). This indicates a systemic error in parsing logic. It is likely that the models extracted a different category (e.g., profit before tax or another related item) and uniformly classified it as Net Result. The fact that the results are identical across models suggests that they rely on similar pattern-recognition algorithms within financial statements, which represents a critical insight for this research.

Gemini Pro demonstrates the highest level of persistence by filling in nearly all cells, but with significant deviations in 2025. It attempts to interpret data even when the probability of accuracy is low. When DeepSeek-V3 is concerned, it exhibits an interesting complete failure in reading Excel data for 2025. The model effectively refuses to process the Excel/CSV format, pointing to more restrictive validation filters or an inability to handle multimodal input in that specific format. GPT-5, on the other hand, produces statistically unacceptable deviations for Operating Income in 2025 (213,926 vs. 90,498). That indicates potential issues with large-number tokenization or incorrect aggregation of subcategories in Excel files.

One of the most critical issues in interpreting balance sheet items is related to "Short-Term Liabilities", which represents a key vulnerability point. For 2024, Gemini Pro reports 19,349 instead of 26,457. In financial analysis, an error in the item of short-term liabilities directly distorts liquidity indicators (Current Ratio, Acid Ratio, Quick Ratio and Net Working Capital). If a researcher were to rely on AI for calculating these ratios, they would obtain a completely misleading picture of the company's solvency and liquidity. This must be clearly identified in the paper as a high operational risk of using AI without human supervision.

Further, Table 2 presents a ratio analysis of liquidity through a comparative overview of results obtained using traditional financial analysis formulas and those generated by modern AI models. The table reveals a systemic deviation in the models of Gemini Pro and DeepSeek V3.

Table 2. Ratio Analysis of Liquidity

RATIO ANALYSIS - LIQUIDITY	**2023 (PDF)**	**Δ%**	**2024 (PDF)**	**Δ%**	**2025 (Excel/CSV)**	**Δ%**
	official		official		unofficial	
Current ratio	**4,29**	**-**	**2,13**	**-**	**2,11**	**-**
Gemini Pro	5,29	+23,31 %	2,91	+36,62 %	2,54	+20,38 %
Chat GPT-5	4,18	-2,56%	2,13	0,00%	2,10	-0,47%
Deep Seek-V3	5,29	+23,31 %	2,91	+36,62 %	2,54	+20,38 %
Quick ratio	**3,87**	**-**	**1,53**	**-**	**0,94**	**-**
Gemini Pro	-	-	-	-	-	-
Chat GPT-5	3,75	-3,10%	1,53	0,00%	0,92	-2,13%
Deep Seek-V3	-	-	-	-	-	-
Acid ratio	**1,08**	**-**	**0,83**	**-**	**0,34**	**-**
Gemini Pro	-	-	-	-	-	-
Chat GPT-5	1,07	-0,93%	0,83	0,00%	0,34	0,00%
Deep Seek-V3	-	-	-	-	-	-
Net working capital	**30.987**	**-**	**29.992**	**-**	**37.664**	**-**
Gemini Pro	-	-	-	-	-	-
Chat GPT-5	-	-	-	-	-	-
Deep Seek-V3	-	-	-	-	-	-

Source: Author's calculations

Specifically, both models exhibit identical deviations ranging from +23% to +36% in the Current Ratio indicator. This strongly suggests that these models share a similar logical error in defining either current assets or short-term liabilities. Such an error is known as "Algorithmically Induced Convergence of Error". Although in Table 2 GPT-5 demonstrated "catastrophic" misreading of revenue in Excel data, it shows surprisingly high precision here in ratio calculations (0.00% error in 2024 across all three categories). This suggests that GPT-5 performs significantly better in formula calculations when provided with correct inputs, but remains vulnerable to "hallucinations" during the data extraction phase from unstructured tables (as observed in the 2025 case).

Another key issue identified in Table 2 is "Incomplete Data Extraction". The fact that Gemini Pro and DeepSeek V3 completely failed in calculating the Quick Ratio and Acid Ratio (most likely due to their inability to isolate "Inventory" from the balance sheet) highlights their inability to perform multi-stage data filtering. This represents a critical barrier to their application in professional financial analysis. Finally, all three models "failed" in calculating Net Working Capital, which serves as clear evidence that AI still struggles to grasp the conceptual distinction between simply retrieving a single value and calculating a residual value (Current Assets − Short-Term Liabilities) in real time. Further, Table 3 presents a ratio analysis of solvency through a comparative overview of results obtained using traditional financial analysis formulas and those generated by modern AI models. A hundred percent correlation between Gemini Pro and DeepSeek V3 is evident in this table.

Table 3. Ratio Analysis of Solvency

RATIO ANALYSIS - SOLVENCY	**12/2023**	**Δ%**	**12/2024**	**Δ%**	**12/2025**	**Δ%**
	official		official		unofficial	
Equity ratio	**77,97%**	**-**	**55,73%**	**-**	**44,81%**	**-**
Gemini Pro	79%	+1,32%	66%	+18,43%	51%	+13,81%
Chat GPT-5	77,3%	-0,86%	55,9%	+0,31%	44,4%	-0,91%
Deep Seek-V3	79%	+1,32%	66%	+18,43%	51%	+13,81%
Gearing - Funded debt to equity ratio	**0,03**	**-**	**0,05**	**-**	**0,30**	**-**
Gemini Pro	-	-	-	-	-	-
Chat GPT-5	1,4%	-53,33%	4,5%	-10,00%	30%	0,00%
Deep Seek-V3	-	-	-	-	-	-
EBITDA / Financial expenses	**110,33**	**-**	**22,81**	**-**	**9,97**	**-**
Gemini Pro	-	-	-	-	-	-
Chat GPT-5	-	-	-	-	-	-
Deep Seek-V3	-	-	-	-	-	-
Long term indebtednes / EBITDA	**0,05**	**-**	**0,57**	**-**	**2,74**	**-**
Gemini Pro	-	-	-	-	-	-
Chat GPT-5	-	-	-	-	-	-
Deep Seek-V3	-	-	-	-	-	-
Total indebtednes / EBITDA	**0,10**	**-**	**0,65**	**-**	**2,80**	**-**
Gemini Pro	-	-	-	-	-	-
Chat GPT-5	-	-	-	-	-	-
Deep Seek-V3	-	-	-	-	-	-
ST indebtedness / Sales	**0,01**	**-**	**0,00**	**-**	**0,00**	**-**
Gemini Pro	-	-	-	-	-	-
Chat GPT-5	-	-	-	-	-	-
Deep Seek-V3	-	-	-	-	-	-

Source: Author's calculations

Their readings of 79%, 66%, and 51% (rounded figures) suggest that these models are not performing precise extraction, but rather approximation. A deviation of 18.43% in 2024 is methodologically unacceptable for rigorous financial analysis. The table also demonstrates that GPT-5 once again emerges as the most reliable tool for ratio analysis. Deviations of less than 1% in the Equity Ratio indicate that the model successfully identifies both “Capital and Reserves” and “Total Assets,” and performs the division correctly. Furthermore, a complete collapse is observed in complex indicators as all models failed in the indicators involving EBITDA. AI models have huge issues when dealing with categories that are not explicitly presented as a single line item in financial statements, but instead require calculations (e.g., Operating Profit + Depreciation + Amortization). Current LLMs are unable to perform derived financial analysis without pre-processed (pre-calculated) data. In the case of GPT-5, a deviation of –53.33% in 2023 is observed for the Gearing. Although the difference is small (0.014 vs. 0.03), the percentage deviation is substantial, indicating that the model confuses long-term liabilities with total liabilities or misinterprets interest liabilities. Further, Table 4 presents a ratio analysis of profitability through a comparative overview of results obtained using traditional financial analysis formulas and modern AI models.

Table 4. Ratio Analysis of Profitability

RATIO ANALYSIS - PROFITABILITY	**2023 (PDF)**	**Δ%**	**2024 (PDF)**	**Δ%**	**2025 (Excel/CSV)**	**Δ%**
	official		official		unofficial	
EBITDA margin	**14,65%**	**-**	**2,53%**	**-**	**4,28%**	**-**
Gemini Pro	-	-	-	-	-	-
Chat GPT-5	-	-	-	-	-	-
Deep Seek-V3	-	-	-	-	-	-
Gross margin	**72,52%**	**-**	**56,69%**	**-**	**66,86%**	**-**
Gemini Pro	-	-	-	-	-	-
Chat GPT-5	-	-	-	-	-	-
Deep Seek-V3	-	-	-	-	-	-
EBIT margin	**12,02%**	**-**	**1,11%**	**-**	**1,65%**	**-**
Gemini Pro	-	-	-		-	-
Chat GPT-5	-	-	-	-	-	-
Deep Seek-V3	-	-	-	-	-	-
Net margin	**10,41%**	**-**	**2,16%**	**-**	**1,96%**	
Gemini Pro	8,74%	-16,04%	2,15%	-0,46%	1,52%	-22,45%
Chat GPT-5	8,6%	-17,39%	2,1%	-2,78%	1,46%	-25,51%
Deep Seek-V3	8,74%	-16,04%	2,15%	-0,46%	1,52%	-22,45%
Rate of return on equity (ROE)	**13,66%**	**-**	**4,84%**	**-**	**3,30%**	**-**
Gemini Pro	-	-	-	-	-	-
Chat GPT-5	17,1%	+25,18%	7,2%	+48,76%	8,94%	+170,91%
Deep Seek-V3	-	-	-	-	-	-
Rate of return on assets (ROA)	**10,28%**	**-**	**2,98%**	**-**	**1,58%**	**-**
Gemini Pro	-	-	-	-	-	-
Chat GPT-5	13,2%	+28,40%	4,0%	+34,23%	3,95%	+150,00%
Deep Seek-V3	-	-	-	-	-	-

Source: Author’s calculations

In Table 4, we have a partial collapse and the inability of analyzing EBITDA, Gross and EBIT margins. Regarding the Net Margin analysis, it is noteworthy that in 2024 (PDF data), all models show negligible error, whereas in 2025 (Excel data), the error rises to over 22%. This confirms earlier findings that AI models, when working with tabular data (Excel/CSV), fail to accurately match Net Income with Operating Revenue when columns are not clearly defined. As for indicators such as ROA and ROE, GPT-5 is the only model that attempted to calculate these values, but with errors reaching up to 170% (ROE in 2025). Why did this occur? ROE requires dividing the items of Net Income by Total Equity. If the model in Table 1 incorrectly read equity, the error in the ratio analysis increases exponentially. This type of error is defined as "Error Propagation in Multi-step Financial Reasoning". The fact that all three models returned empty fields ("–") for EBITDA and Gross margins across all years indicates that LLMs are unable to calculate "Cost of Goods Sold" (COGS) or EBITDA from raw financial statements when these items are not explicitly stated.

This table finalizes the argument that, while AI is capable of extracting data with errors, it is fundamentally incapable of performing independent financial synthesis. The more complex the ratio (i.e., the more data sources it requires), the higher the probability of inaccuracy or complete failure (Data Omission).

4. Discussion

Empirical data obtained from the research results demystify the infallibility of AI models in financial reporting. An inverse correlation has been observed between format complexity (Excel) and extraction accuracy. Particularly concerning is the occurrence of inaccuracies where different models generate identical erroneous values, implying weaknesses in the mapping of financial statement items. These findings impose the necessity of introducing "Human-in-the-loop" protocols as an imperative in both academic and professional financial research. The paper showed that if the data is not structured correctly that generic LLMs models make mistakes, which is somewhat consistent with the research of Liu (2023).

The results presented in the tables refute the myth of the infallibility of large language models (LLMs) in the domain of precise financial analysis. The study identified fundamental structural limitations that prevent the autonomous application of AI in corporate finance without rigorous human verification. Accordingly, hypothesis H1, which states: "When independently analyzing complex financial statements, without human intervention, generic LLM models (GPT-5, Gemini Pro, and DeepSeek V3) have no errors or limitations," is fully rejected.

Furthermore, it has been empirically demonstrated that the transition from visually structured PDF reports to tabular Excel/CSV formats does not facilitate, but rather impairs the analytical precision of the models. While PDFs enable the use of spatial reasoning, Excel files without clearly defined metadata lead to "drifting". A deviation of 136.39% in operating revenues in 2025 (ChatGPT-5) represents a statistical absurdity, indicating the model's complete inability to recognize the hierarchy of financial statement items in a non-standardized tabular environment. At the center of the research on the necessity of integrating AI and financial analysis is the problem of reliability which is fully consistent with the results of the research by Brown (2024).

The study reveals a phenomenon referred to as the "Exponential correlation of inaccuracies". A minor error in extracting a base financial item results in catastrophic distortions in solvency, liquidity, and profitability ratios. Deviations in the Equity ratio (up to 18.43%) and ROE (up to 170.91%) indicate that LLMs lack an internal "accounting logic" that would signal when obtained results are economically implausible. Instead, they operate on the principle of statistical probabilities of character sequences rather than on the principle of balance sheet equilibrium. The observed deviations in the research indicate that the research is fully consistent with the findings of Johnson (2024), only with even higher percentages of deviation.

The scientific contribution of this paper lies in identifying the imperfections of AI models when dealing with derived categories such as EBITDA or gross margin. The complete absence of results across all tested models for these items confirms that artificial intelligence, despite its generative capabilities, remains "blind" to concepts that require prior data transformation and a deep understanding of accounting standards. The above indicates that our research results are fully consistent with the results obtained by the author Zhang (2023).

Consequently, hypothesis H2, which states: "There is a necessity for integrating AI and financial analysis so that the synergy between humans and AI results in superior analytical outcomes, enabling faster decision-making with a significantly reduced error rate," is fully confirmed.

Artificial intelligence in its current form (GPT-5, Gemini Pro, DeepSeek V3) can serve only as a tool for accelerated initial scanning, but not as a substitute for expert judgment. The introduction of AI into processes of strategic investment decision-making, without accompanying "Human-in-the-loop" protocols, represents a high level of operational and reputational risk. The necessity of integration is not reflected in replacing analysts, but in drastically accelerating data processing and decision-making. AI integration enables analysts to compare hundreds of reports in seconds, identify patterns that deviate from historical averages, and test various scenarios, while final verification and strategic judgment remain the responsibility of humans.

This paper establishes the foundation for a new paradigm of "hybrid intelligence," in which humans remain the ultimate arbiters of truth, while AI functions merely as an accelerated (yet currently unreliable) processor of raw information. The topic of this research paper is highly relevant, as it demonstrates that not everything can be solved with a single click of artificial intelligence. At the same time, it is modern, as it incorporates the use of current AI tools and proves that human expertise remains irreplaceable, but must be technologically enhanced.

REFERENCES

1. Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance*, 23(4), 589–609. https://doi.org/10.1111/j.1540-6261.1968.tb00843.x
2. Brown, A. (2024). *Analytical hallucinations in AI-driven financial reporting*. Journal of Digital Finance & AI, 12(2), 45–58.
3. Cao, L., et al. (2022). *Modern economic conditions and the rigidity of classical financial models*. International Journal of Financial Studies, 10(1), 112–130.
4. DeepSeek-AI. (2024/2025). *DeepSeek-V3 Technical Report*. DeepSeek, (tool for analysis – Access date – April 2026.).
5. Google. (2024). *Gemini Pro: A new family of highly capable multimodal models*. Google DeepMind, (tool for analysis – Access date – April 2026.).
6. Hsu, Y. (2024). *Synthesis of large-scale financial documentation using generative AI*. Tech-Finance Review, 5(3), 201–215.
7. Johnson, R. (2024). *Error rates in autonomous AI interpretation of complex accounting items*. Accounting & Technology Quarterly, 8(1), 14–29.

8. Liu, S. (2023). *Superiority of machine learning in structured financial data analysis*. Journal of Data Science and Analytics, 15(4), 88–104.
9. Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The Journal of Finance*, 66(1), 35–65. https://doi.org/10.1111/j.1540-6261.2010.01625.x
10. OpenAI. (2024/2025). *GPT-5 System Card and Technical Documentation*. OpenAI, (tool for analysis – Access date – April 2026.).
11. Smith, J., & Doe, J. (2025). *The augmentation model: Human-in-the-loop in financial forecasting*. Future of Finance Press.
12. Wang, X., et al. (2024). *Master Prompting: New essential skills for financial analysts in the AI era*. Global Finance Journal, 31(2), 77–92.
13. Zhang, L. (2023). *The gap between linguistic superiority and accounting logic in LLMs*. AI & Business Research, 19(3), 150–166.