

LARGE LANGUAGE MODELS EMPOWERING COMPLIANCE CHECKS AND REPORT GENERATION IN AUDITING

ZhiWen Gan

School of Business Administration, Baise University, Baise 533000, Guangxi, China.

Corresponding Email: 2201776207@qq.com

Abstract: This study aims to explore the potential application of large language models (LLMs) in compliance checks and report generation in auditing. Through literature analysis and theoretical discussion, this paper examines the advantages of LLMs in handling unstructured data and automatically generating audit reports. The research findings suggest that LLMs, with their powerful text processing and generation capabilities, can automatically identify potential compliance risks and generate high-quality audit reports, significantly enhancing audit efficiency. Meanwhile, this study also highlights that challenges such as model interpretability and data security remain major obstacles in their application. The study concludes that LLMs will play a critical role in future intelligent auditing processes, providing technical support to improve audit work efficiency.

Keywords: Large language models; Compliance checks; Audit report generation; Intelligent auditing; Unstructured data processing; Natural language processing

1 INTRODUCTION

Auditing, as a core function of corporate financial management and compliance oversight, plays a crucial role. With the continuous development of the global economy, the complexity of business operations has increased, and the challenges faced by auditors have grown more severe. Modern auditing not only ensures the accuracy of financial statements but also verifies whether companies comply with relevant laws, regulations, industry standards, and internal control requirements. This has made the tasks of auditors more complex, especially in companies involved in cross-border operations or multi-industry activities, where compliance checks have become more burdensome [1]. Moreover, the diversity of audit data and the variety of information sources, especially the surge of unstructured data such as legal contracts, policy documents, emails, and meeting minutes, have significantly increased the workload of auditors [2]. These challenges have rendered traditional auditing tools and methods inefficient, creating an urgent need for new technologies to improve the accuracy and efficiency of audits.

In this context, the development of artificial intelligence (AI) technologies has brought new opportunities to the field of auditing. Particularly, the emergence of large language models (LLMs) has transformed the way we process and understand textual data. LLMs can capture complex linguistic patterns through deep learning from vast amounts of textual data, enabling precise understanding and generation of natural language. Their powerful language processing capabilities have demonstrated unique advantages in handling complex and unstructured data, particularly in fields that heavily rely on textual information such as finance, law, and auditing [3].

In auditing, compliance checks and report generation have always been key tasks. Compliance checks require auditors to quickly identify non-compliant content from large amounts of textual data, which is time-consuming and prone to human error. The introduction of LLMs can help automate the processing and analysis of these complex texts, quickly identifying potential compliance issues. For example, by analyzing corporate contracts, legal documents, and policies, the model can determine whether any provisions violate relevant laws and regulations and promptly alert auditors to take action. At the same time, in report generation, LLMs can automatically summarize and analyze audit data to generate high-quality audit reports, providing decision-makers with clear compliance analyses and reducing auditors' workloads.

Furthermore, with the increasingly stringent global regulatory environment, compliance auditing has become a critical component for companies to maintain competitiveness within complex legal frameworks. The growing regulatory pressure has driven auditing work toward intelligent transformation, and the application of LLMs is a key force behind this transformation. Utilizing LLMs not only effectively reduces audit costs but also enhances the timeliness and accuracy of audits by automating audit processes and reducing audit risks.

Therefore, this paper focuses on exploring the potential application of LLMs in compliance checks and report generation in auditing, discussing their application scenarios and advantages in practice. By analyzing their working principles and application effects, this paper aims to construct an intelligent auditing framework based on LLMs, providing theoretical support and practical guidance for the intelligent development of future audit processes.

2 LITERATURE REVIEW

With the rapid development of big data and artificial intelligence technologies, research in the field of auditing has gradually moved towards intelligence and automation. Traditional auditing tools and methods can no longer effectively address the complexities of today's audit environment, especially when handling vast amounts of unstructured data. Auditors often rely on manual processing, which is inefficient and prone to errors. In this context, an increasing number of studies have started exploring how technological solutions can improve the efficiency and accuracy of audits, with the application of large language models (LLMs) being particularly noteworthy.

First, existing studies have explored the initial application of natural language processing (NLP) technology in the field of auditing. NLP can help auditors extract useful information from large volumes of textual data, automate the processing of audit evidence, and generate preliminary audit reports. For example, some studies have demonstrated the use of NLP in contract analysis, compliance review, and financial report generation, highlighting its advantages in handling complex textual data [4]. These studies lay the foundation for the further application of LLMs, as LLMs can handle more complex language tasks than traditional NLP models, especially when dealing with large-scale unstructured data [5-6].

Second, the rise of LLMs has brought revolutionary advancements to text processing technologies. Compared to traditional auditing tools, LLMs can learn from vast datasets through pre-training, extract complex linguistic patterns, and understand context and semantics to perform more intricate language tasks [7]. Research shows that LLMs exhibit strong generalization capabilities across various domains, especially in fields such as finance and law, which heavily rely on textual information [8]. For example, in the legal field, LLMs have been used to analyze legal documents, automatically generate legal opinions, and assist in contract reviews, showcasing their powerful text understanding and generation capabilities [9]. These successful cases provide a reference for their application in compliance checks within auditing.

Additionally, some studies specifically explore the prospects of applying LLMs in compliance checks in auditing. LLMs can quickly and accurately extract key information from large volumes of text and identify potential compliance risks [10]. The core advantage of this technology lies in its ability to generate contextually appropriate responses automatically, helping auditors quickly identify content that is inconsistent with regulations or internal control standards. For example, in financial audits, models can analyze company contracts and policy documents to detect potential legal loopholes or violations [11]. This automated compliance checking mechanism not only improves audit efficiency but also significantly reduces subjective biases in manual auditing.

Finally, although the application of LLMs in auditing holds significant promise, some scholars have raised potential challenges and issues. For instance, the interpretability and transparency of the models are important areas of concern. In a rigorous field such as auditing, the decision-making processes of the models must be transparent, enabling auditors to understand the basis for the model's judgments [12-13]. Moreover, data privacy and security are critical concerns in the application of LLMs, particularly when handling sensitive corporate data. Ensuring data security and privacy protection will be a key focus of future research [14].

In summary, this literature review highlights the initial applications and development trends of LLMs and related technologies in the field of auditing. Through a review of existing research, it is evident that LLMs have significant advantages in handling complex textual data and automatically generating audit reports. However, challenges such as model interpretability and data security remain, and future research needs to explore how to overcome these technical barriers.

3 THEORETICAL ANALYSIS

The ability of large language models (LLMs) to demonstrate great potential in the field of auditing, particularly in compliance checks and report generation, is primarily attributed to their powerful text processing and language generation capabilities. This section will analyze in detail the specific roles LLMs play in the auditing process, explore their application scenarios and advantages, and discuss the potential challenges they may face.

3.1 Working Mechanism of Large Language Models

LLMs are based on deep learning and pre-training techniques, learning complex patterns and semantic relationships in language through exposure to vast amounts of textual data. Pre-trained LLMs can understand complex contexts and generate text that fits the given scenario. In the context of compliance checks in auditing, the models can quickly analyze and process large volumes of text data, such as policy documents, legal contracts, and internal reports. LLMs are capable of identifying potential non-compliance issues and can even assess the severity of compliance risks based on contextual associations, providing auditors with precise decision support.

For example, when auditing a company's financial report, LLMs can automatically scan the company's policy documents to identify provisions that may conflict with industry regulations or internal control standards, and flag potential risks. This automated auditing process greatly enhances efficiency, particularly when dealing with unstructured data, where the models' performance is especially noteworthy. Using natural language processing (NLP) technology, auditors can quickly extract key information from large volumes of unstructured text without the need to manually check each document.

3.2 Application Scenario Analysis

The specific applications of LLMs in compliance checks and report generation within auditing can be categorized into the following areas:

First, compliance review automation. LLMs can automatically extract information from unstructured text, such as internal corporate documents and external regulatory documents, to detect potential compliance issues. By comparing the extracted information with industry regulations and company standards, the model can identify compliance risks at an early stage, helping companies make timely corrections. This application is especially beneficial for large

multinational companies, as they are often subject to multiple jurisdictions' laws and regulations, making manual compliance reviews extremely labor-intensive and prone to oversight.

Second, automatic audit report generation. The process of generating audit reports involves organizing and analyzing large amounts of data, which usually requires auditors to summarize various findings into written reports. LLMs can analyze the data generated during the audit and automatically produce structured audit reports. The models can not only summarize audit findings but also generate explanatory notes and recommendations based on context. This application can significantly reduce the time auditors spend on report writing while ensuring the consistency and professionalism of the reports. Just as hybrid teaching models improve the efficiency of medical interns, the large-scale application of intelligent auditing tools can greatly enhance auditors' work efficiency through training and support [15].

Finally, risk assessment and decision support. LLMs can also be used in auditing for risk assessment. By analyzing historical and current audit data, the models can identify high-risk areas or transactions and provide both quantitative and qualitative analysis support. For example, the model can analyze a company's past compliance history, financial data, and the relevance of external regulations to prioritize high-risk matters for auditors. This data-driven audit process improves the scientific and accuracy of audit decision-making.

3.3 Technical Advantages

Compared to traditional auditing methods, LLMs offer significant advantages in the following aspects:

First, the ability to handle unstructured data. Traditional auditing tools are often limited to processing structured financial data and struggle with large volumes of unstructured data, such as contract texts, emails, and meeting notes. LLMs, with their deep understanding of text, can easily handle such data, quickly extracting key information and greatly reducing the manual workload of auditors. Just as comparative studies of different treatment methods in the medical field have shown, intelligent auditing tools excel in automating complex compliance review tasks, particularly in processing unstructured data [16].

Second, flexibility in language understanding and generation. LLMs can not only comprehend complex linguistic structures but also generate logical and coherent text. This makes them uniquely advantageous in generating audit reports and summarizing documents. Auditors can use the model to quickly generate high-quality reports, ensuring accuracy and consistency in the content.

Finally, broad adaptability and scalability. LLMs can be fine-tuned and trained to meet the auditing needs of different industries and companies. For example, companies can customize the models to meet their specific compliance requirements, allowing the model to exhibit higher professionalism and accuracy in the auditing process.

3.4 Challenges and Future Research Directions

Despite the significant potential of LLMs in auditing, they also face challenges in practical application. First, the explainability of the models remains a key issue. Auditing requires a high level of transparency and rigor, but the decision-making process of LLMs is often difficult to fully understand, making it challenging for auditors to verify the basis of the model's judgments. Similar to the requirements for karyotype analysis in the field of biology, the application of LLMs in auditing also demands high explainability and accuracy in model outputs to ensure that auditors can fully comprehend and verify the model's judgments [17].

Second, data privacy and security issues are critical considerations. Auditing involves handling large amounts of sensitive corporate data, and ensuring data security when using LLMs is an urgent technical problem to be solved. Similar to the Rosenthal effect in nursing interventions, the use of LLMs in auditing must also guard against potential problems caused by biases in training data [18]. Protecting privacy and ensuring ethical compliance are crucial, just as they are in teaching reform focusing on morality and ethics [19].

In summary, LLMs have great potential in compliance checks and report generation in auditing, capable of improving audit efficiency through automation and reducing human errors. However, future research must focus on enhancing model explainability, strengthening data security, and optimizing model performance in complex auditing scenarios.

4 CONCLUSION

From both theoretical and practical perspectives, this paper discusses the potential application value of large language models (LLMs) in the field of auditing, particularly in compliance checks and report generation. By summarizing the working principles of LLMs, their specific application scenarios in auditing, and the optimization of audit processes, the following conclusions can be drawn:

First, LLMs possess powerful text processing and generation capabilities, making them highly effective in handling large volumes of unstructured data during audits. Compared with traditional auditing tools, LLMs can quickly analyze complex documents such as contracts, policy files, and financial reports, and efficiently identify potential compliance issues. Their automation capabilities significantly reduce the need for manual intervention, improve audit efficiency, and reduce the workload of auditors, especially when dealing with heavy data processing tasks, ultimately streamlining the entire auditing process.

Second, LLMs show significant advantages in audit report generation. By analyzing the data produced during the audit process, these models can automatically summarize key audit findings and generate structured, coherent audit reports

based on the context. This not only improves the accuracy and consistency of reports but also shortens the time needed for report creation, allowing auditors to focus more on high-value judgment and decision-making tasks.

However, despite the great potential of LLMs in auditing, several challenges still remain in practical application. First, the explainability and transparency of the models continue to be critical obstacles. In a field as rigorous as auditing, auditors need to clearly understand the basis for the model's judgments to ensure that its conclusions are reliable. Additionally, data privacy and security are critical concerns, especially when handling sensitive corporate data. Ensuring the security of data during the use of LLMs will be a key issue that needs to be addressed in future applications.

Future research can focus on the following directions. First, efforts should be made to enhance the explainability of LLMs, enabling auditors to better understand and trust the model's outputs. Second, ensuring data security when processing sensitive data with LLMs is a crucial area for further research. Moreover, as LLMs are applied in more auditing scenarios, research should also explore how these models can be integrated with other auditing technologies to achieve more intelligent and automated audit processes.

In summary, LLMs provide significant technical support for the intelligence and automation of audit processes, particularly in the areas of compliance checks and report generation. As the technology continues to develop and application challenges are gradually overcome, LLMs are expected to play an increasingly important role in the field of auditing, becoming essential tools for auditors to work more efficiently.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

- [1] Luo H. Application and influencing factors of audit technology methods innovation in the digital intelligence era. *Journal of Xihua University: Philosophy and Social Sciences Edition*, 2023, 42(4): 29-37.
- [2] Appelbaum D, Showalter DS, Sun T, et al. A framework for auditor data literacy: A normative position. *Accounting Horizons*, 2021, 35(2): 5-25.
- [3] Liu T, Zhang SR, Li ZH, et al. Intelligent classification method of water project inspection text based on character-level CNN. *Journal of Hydroelectric Engineering*, 2021, 40(6): 10.
- [4] Schumann G, Gómez JM. Natural Language Processing in Internal Auditing—a Structured Literature Review. *AMCIS*, 2021.
- [5] Bin Abdullah MR, Iqbal K. A review of intelligent document processing applications across diverse industries. *Journal of Artificial Intelligence and Machine Learning in Management*, 2022, 6(2): 29-42.
- [6] Cejas OA, Azeem MI, Abualhaija S, et al. NLP-based automated compliance checking of data processing agreements against GDPR. *IEEE Transactions on Software Engineering*, 2023, 49(9): 4282-4303.
- [7] Hassan FU, Le T, Lv X. Addressing legal and contractual matters in construction using natural language processing: A critical review. *Journal of Construction Engineering and Management*, 2021, 147(9): 03121004.
- [8] Fantoni G, Coli E, Chiarello F, et al. Text mining tool for translating terms of contract into technical specifications: Development and application in the railway sector. *Computers in Industry*, 2021, 124: 103357.
- [9] Antos A, Nadhamuni N. Practical guide to artificial intelligence and contract review. In: *Research Handbook on Big Data Law*. Edward Elgar Publishing, 2021, 467-481.
- [10] Yan JZ, He YX, Luo ZY, et al. Potential applications and challenges of generative large language models in the medical field. *Journal of Medical Informatics*, 2023, 44(9): 23-31.
- [11] Dong WL, Liu Z, Liu K, et al. A review of smart contract vulnerability detection technologies. *Journal of Software*, 2023, 35(1): 38-62.
- [12] Thunki P, Reddy SRB, Raparathi M, et al. Explainable AI in Data Science: Enhancing Model Interpretability and Transparency. *African Journal of Artificial Intelligence and Sustainable Development*, 2021, 1(1): 1-8.
- [13] Delcaillau D, Ly A, Papp A, et al. Model transparency and interpretability: Survey and application to the insurance industry. *European Actuarial Journal*, 2022, 12(2): 443-484.
- [14] Maruthi S, Dodda SB, Yellu RR, et al. Language Model Interpretability—Explainable AI Methods: Exploring explainable AI methods for interpreting and explaining the decisions made by language models to enhance transparency and trustworthiness. *Australian Journal of Machine Learning Research & Applications*, 2022, 2(2): 1-9.
- [15] Wu PX, Lan ZY, Huang W, et al. Application of blended teaching mode based on MOOC in medical undergraduate internships. *Journal of Youjiang Medical University for Nationalities*, 2020, (1): 119-122.
- [16] Rao ZF, Chen JY, Zhao LY, et al. Comparison of ultrasound-guided microwave ablation and laparoscopic resection in the treatment of benign thyroid nodules. *Journal of Youjiang Medical University for Nationalities*, 2018, (6): 583-585.
- [17] Yang MX, Niu SR. Karyotype analysis of peripheral blood lymphocyte culture and G-banding in 5309 couples with adverse pregnancy histories. *Journal of Youjiang Medical University for Nationalities*, 2019, (5): 520-522.

- [18] Cui XL, Yang ML. The effect of nursing intervention based on the Rosenthal effect on the stress response and emotional state of children with scoliosis. *Journal of Youjiang Medical University for Nationalities*. 2019, (6): 713-715.
- [19] Pang XR, Liu F, Li L, et al. Preliminary exploration of teaching reform of "General Medicine" based on the concept of "Curriculum Ideology and Politics". *Journal of Youjiang Medical University for Nationalities*, 2020, (6): 806-809.