INTERNATIONAL JOURNAL OF PUBLIC FINANCE AND ACCOUNTABILITY

NAWALA

Volume 3, Number 2, *2024*

Machine Learning in Government Auditing: A Systematic Literature Review

Hadaina Rusyda [1*]

[1] Universitas Diponegoro, Semarang, Indonesia

Abstract

Article history:
Received: July 6, 2024
Revised: August 22, 2024
Accepted: October 28, 2024
Published: December 30, 2024

Keywords:
Fraud Detection, Government Auditing, Machine Learning, Public Sector, Transparency.

Identifier:
Nawala
Page: 71-79
https://nawala.io/index.php/ijpfa

The rapid advancement of Machine Learning (ML) has begun to transform government auditing by enabling full-population testing, anomaly detection, and predictive risk assessment. This study employs a systematic literature review (SLR) of peer-reviewed articles published between 2016 and 2023 to examine the applications, benefits, and challenges of ML in public sector auditing. Findings indicate that ML enhances audit efficiency, supports data-driven decision-making, and improves fraud detection. However, technical limitations, data quality issues, and the "black box" nature of certain algorithms pose significant challenges, raising concerns about transparency and accountability. Ethical and governance considerations are critical for ensuring legitimacy and public trust, while the fragmented nature of existing research highlights the need for standardized evaluation frameworks and interdisciplinary expertise. Overall, the study underscores the transformative potential of ML in government auditing while emphasizing the importance of responsible and structured implementation.

*Corresponding author:
(Hadaina Rusyda)

1. Introduction

Over the past decade, the rapid advancement of artificial intelligence (AI), particularly machine learning (ML), has begun to transform the practice of government auditing. In the public sector where transaction volumes, administrative documents, and program data are vast ML enables auditors to conduct full-population testing, detect anomalies that might escape traditional sampling, and generate predictive risk assessments for more targeted oversight (Huang et al., 2022; Kokina & Davenport, 2017). Such capabilities promise to enhance efficiency, improve fraud detection, and support evidence-based decision-making in audit planning. As government agencies increasingly embrace data-driven practices, ML stands out as a critical tool for modernizing Supreme Audit Institutions (SAIs) and strengthening public accountability.

Despite these benefits, integrating ML into government auditing poses substantial challenges. Technical issues, such as data quality, feature engineering, and model validation, can limit the effectiveness of ML applications (Fedyk et al., 2022). More importantly, the "black-box" nature of some algorithms raises transparency and explainability concerns, making it difficult for auditors to justify their findings to policymakers and the public (Munoko et al., 2020). In the public sector context where legitimacy depends on impartiality and trust these ethical and governance concerns are particularly salient (Mökander & Floridi, 2021). Addressing such challenges requires not only technical expertise but also the establishment of clear standards, audit frameworks for algorithms, and mechanisms for accountability in AI-assisted auditing.

Given the growing body of research over the last decade, scholarly interest in ML applications for auditing is evident, yet the literature remains fragmented. Studies vary widely in their methodological approaches, audit domains, and types of ML techniques used from anomaly detection and predictive risk scoring to natural language processing for document analysis. This diversity, while enriching, makes it difficult to form a cohesive understanding of the field's current state. A systematic literature review (SLR) is therefore essential to synthesize existing evidence, identify practical and theoretical contributions, and highlight gaps that warrant further exploration. Such a synthesis can provide valuable insights for both researchers and practitioners, guiding the responsible and effective integration of ML into government auditing.

2. Literature Review

In recent years, the integration of Machine Learning (ML) into public sector auditing has garnered significant attention due to its potential to enhance audit processes. ML techniques enable auditors to analyze large datasets efficiently, identify anomalies, and predict risks, thereby improving the effectiveness and efficiency of audits. For instance, studies have demonstrated that ML algorithms can detect fraudulent activities by analyzing patterns and anomalies in financial data (Straub et al., 2022). These capabilities promise to enhance efficiency, improve fraud detection, and support evidence-based decision-making in audit planning.

Despite the promising capabilities of ML, its adoption in government auditing presents several challenges. Technical issues such as data quality, feature engineering,

and model validation can limit the effectiveness of ML applications (Green, 2021). Moreover, the "black-box" nature of some ML algorithms raises concerns about transparency and explainability, making it difficult for auditors to justify their findings to policymakers and the public (Costanza-Chock et al., 2023). These challenges underscore the need for developing clear standards and frameworks to guide the integration of ML into auditing practices (Green, 2021).

Furthermore, the literature highlights the importance of a systematic approach to evaluating ML applications in auditing. Systematic literature reviews (SLRs) have been conducted to synthesize existing research, identify trends, and assess the effectiveness of various ML techniques in fraud detection and risk assessment (Anastasopoulos & Whitford, 2018). Such reviews provide valuable insights into the current state of ML in auditing and offer directions for future research and practice (Anastasopoulos & Whitford, 2018).

3. Methods

This study employs a Systematic Literature Review (SLR) to examine the application of Machine Learning in government auditing. The SLR approach enables a comprehensive and structured analysis of existing research, providing insights into key trends, challenges, and gaps in the literature. Relevant studies were identified through searches in academic databases using keywords related to machine learning, government auditing, public sector audit, and fraud detection. Only peer-reviewed articles published between 2016 and 2023 were considered to ensure that the review reflects the most recent developments in the field.

Inclusion criteria focused on studies that addressed ML applications in the public sector or government auditing, discussed technical or ethical challenges, or evaluated the impact of ML on audit efficiency, risk assessment, and fraud detection. Studies that concentrated on ML applications outside the public sector or were not published in English were excluded. Selected studies were analyzed to extract key information such as the type of ML techniques, audit domains, dataset sizes, methodological approaches, reported benefits, and identified challenges. Ethical, governance, and transparency considerations were also examined.

The extracted data were synthesized qualitatively to identify patterns, trends, and areas requiring further research. This approach provides a holistic overview of current ML applications in government auditing and offers guidance for the effective and responsible integration of ML in public sector audit practice.

4. Results and Discussion

The systematic literature review revealed that Machine Learning (ML) has significantly impacted government auditing by enhancing efficiency, improving fraud detection, and supporting data-driven decision-making. Across the studies analyzed, ML techniques such as anomaly detection, predictive risk scoring, and natural language processing were frequently employed to handle large datasets in public sector audits (Huang et al., 2022; Straub et al., 2022). These methods allow auditors to perform full-population testing, identify irregular patterns, and prioritize high-risk areas, thereby improving the effectiveness and timeliness of audit processes.

Despite these advantages, several challenges emerged consistently in the literature. Technical limitations, including data quality issues, feature engineering difficulties, and model validation constraints, were identified as barriers to optimal ML adoption in auditing (Green, 2021; Fedyk et al., 2022). Additionally, the "black-box" nature of certain algorithms raised concerns about transparency and explainability, complicating auditors' ability to justify their findings to policymakers and stakeholders (Costanza-Chock et al., 2023; Munoko et al., 2020). These findings highlight the tension between leveraging sophisticated ML tools and maintaining accountability and trust in the public sector.

Ethical and governance considerations also surfaced as critical themes. Studies emphasized that integrating ML into auditing requires clear standards, accountability mechanisms, and frameworks for algorithmic oversight to ensure that audits remain impartial and legitimate (Mökander & Floridi, 2021; Green, 2021). Moreover, the review indicated that interdisciplinary expertise combining technical ML knowledge with public sector governance understanding is essential for effective implementation.

Another notable trend was the fragmented nature of the literature. Research varies widely in methodological approaches, types of ML techniques, and audit domains, making it challenging to consolidate best practices or draw generalized conclusions (Anastasopoulos & Whitford, 2018). However, this diversity also points to rich opportunities for future research, particularly in exploring standardized evaluation frameworks for ML-based auditing, improving algorithm explainability, and assessing long-term impacts on public sector accountability.

Overall, the findings suggest that while ML offers transformative potential for government auditing, its effective deployment hinges on addressing technical, ethical, and governance challenges. Policymakers and auditing institutions are encouraged to adopt structured implementation frameworks, promote transparency, and invest in auditor training to maximize the benefits of ML while mitigating associated risks.

5. Conclusion

This study highlights the transformative potential of Machine Learning (ML) in government auditing, particularly in enhancing efficiency, improving fraud detection, and supporting data driven decision making. ML techniques such as anomaly detection, predictive risk scoring, and natural language processing enable auditors to analyze large datasets and identify high risk areas more effectively. However, the integration of ML into public sector auditing also presents significant challenges, including technical limitations, data quality issues, and the "black-box" nature of certain algorithms, which raise concerns about transparency and accountability.

Ethical and governance considerations are crucial for ensuring that ML assisted audits maintain legitimacy and public trust. Establishing clear standards, accountability frameworks, and auditor training are essential steps for mitigating these challenges. Additionally, the fragmented nature of existing research indicates a need for further studies on standardized evaluation frameworks, explainability, and long-term impacts of ML in auditing. Overall, while ML offers substantial benefits

for government auditing, careful and responsible implementation is necessary to balance innovation with accountability and transparency.

References

Anastasopoulos, L. J., & Whitford, A. B. (2018). Machine learning for public administration research, with application to organizational reputation. *arXiv*.

Costanza-Chock, S., Harvey, E., Raji, I. D., Czernuszenko, M., & Buolamwini, J. (2023). Who audits the auditors? Recommendations from a field scan of the algorithmic auditing ecosystem. *arXiv*.

Fedyk, A., Hodson, J., Khimich, N. V., & Fedyk, T. (2022). Is artificial intelligence improving the audit process? *Review of Accounting Studies, 27*(3), 938–985.

Green, B. (2021). The flaws of policies requiring human oversight of government algorithms. *arXiv*.

Huang, F., No, W. G., Vasarhelyi, M. A., & Yan, Z. (2022). Audit data analytics, machine learning, and full population testing. *Journal of Finance and Data Science, 8*, 138–144.

Kokina, J., & Davenport, T. H. (2017). The emergence of artificial intelligence: How automation is changing auditing. *Journal of Emerging Technologies in Accounting, 14*(1), 115–122.

Mökander, J., & Floridi, L. (2021). Ethics-based auditing to develop trustworthy AI. *Minds and Machines, 31*(2), 323–327.

Munoko, I., Brown-Liburd, H. L., & Vasarhelyi, M. A. (2020). The ethical implications of using artificial intelligence in auditing. *Journal of Business Ethics, 167*(2), 209–234.

Straub, V. J., Morgan, D., Bright, J., & Margetts, H. (2022). Artificial intelligence in government: Concepts, standards, and a unified framework. *arXiv*.