International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 10 October 2026
Page No: 5958 - 5967
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0764

AI-Driven Risk-Based Auditing Framework for Automated Anomaly Detection and Financial Misstatement Identification

Sachin Kumar Gupta[1]
[1]*University of the Cumberlands, Kentucky*
Email id: *sachinugupta1@gmail.com*[1]

Abstract
The adoption of AI as an enabler for the risk-based audit approach is gaining traction as manual sampling and manually defined tests may overlook financial anomalies that occur at low frequencies and high impact. In this review we critically analyse peer-reviewed studies linking audit data analytics, artificial intelligence, machine learning, automated anomaly detection and financial misstatement identification. There are three complementary research streams identified in the literature that constitute topics for further research: audit-risk and evidence architectures, which help inform where analytics should fit into the engagement; data analytic approaches, which build on audit testing to move toward full-population transaction testing; and supervised or hybrid machine-learning models that aim to distinguish materially misstated or fraudulent financial reporting. There are reported significant improvements in coverage, prioritization and predictive discrimination, and limited evidence of direct integration to assertion-level audit risk. Issues that persist include: the difficulty of learning from rare events; the burden of false positives; the lack of data provenance; the lack of transparency of models; inconsistent validation; the challenge to organize the validation; and the lack of connection between transaction anomalies and material financial-statement effects. Calibrated anomaly scoring, contextual audit evidence, interpretable escalation rules and explicit misstatement-risk mapping are thus necessary to implement a robust risk-based architecture. The field is very promising, but the end-to-end frameworks, which are operationally proven, are still relatively immature.

Keywords: *AI-driven auditing; anomaly detection; financial misstatement; machine learning; risk-based auditing*

1. Introduction
Many of today's financial audits take place in settings that feature complex business operations, varying enterprise data, automated accounting methods, and frequent control changes. In spite of its value, conventional audit sampling is not so compelling an economic choice as computational methods can sample large transaction populations and focus on observing those transactions that are estimated to pose the greatest risk. Cao, Chychyla and Stewart [1] suggest that Big Data analytics can significantly change financial-statement auditing by providing greater breadth of coverage and relationships that are often not captured by traditional tests. Brown-Liburd, Issa and Lombardi [2] demonstrate, however, that while expanded analytical capacity does not necessarily imply improved audit quality, information volume, complexity and ambiguity can lead to an erroneous judgment. It is not, therefore, just a matter of automated detection, but one of the development of an AI-based risk-based auditing framework that can convert anomalous signals identified by algorithms to audit relevant evaluations of potential financial misstatement. Artificial Intelligence brings in computational mechanisms to learn nonlinear patterns, rank abnormal transactions, fuse together structured and unstructured indicators and continually tune the detection models. Issa, Sun, and Vasarhelyi [3] view AI as a tool for automating repetitive audit work and augmenting professional judgment in tasks that call for pattern recognition and exception screening. Appelbaum, Kogan, and Vasarhelyi [4] build on this notion, identifying advanced analytics as a component of the modern audit engagement in which the acquisition, validation, analytical work, assessment of evidence, and interpretation of the evidence by the auditor are interdependent. This type of research is important because it makes a distinction between accuracy of prediction and usefulness in the audit. A model can be

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 10 October 2026
Page No: 5958 - 5967
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0764

a good classification model but be ineffective for risk-based auditing if outputs aren't traceable to assertions, control failures, materiality or plausible misstatement mechanisms. The research therefore is interdisciplinary. While auditing research focuses on the reliability of evidence, professional judgment, risk assessment, and the implementation of the engagement, information-systems and machine-learning research focuses on data preparation, feature construction, classification, anomaly scoring, and predictive validation. Financial-fraud research methods that directly impact material misstatements, such as those that address rare events, add important tools to the detection arsenal. There are gaps in integration as complementary contributions are not fully applied. Some audit analytics studies also end at the point of anomaly identification or feasibility to implement, while some fraud classification studies examine predictive discrimination without modelling audit assertions, tolerable misstatement, control reliance, or engagement workflow. It is thus necessary to conduct a focused review to evaluate how these research streams can be leveraged for an AI-based risk-based auditing framework for automated detection of anomalies and misstatements in financial statements. This review focuses on concepts and methodology, outcomes, implementation challenges, and unaddressed connections between anomalous accounting activity and statement-level misstatement risk. A focus is placed on the ability of current methods to prioritize risks over generating exceptions and whether predictive models provide evidence that is appropriate for professional judgment and on whether validation designs reflect the actual conditions at audit engagements. The assessment reveals an important transition space between isolated analytics and auditable, explainable, risk calibrated systems that relate signals at the transaction level to assertions at the financial-statement level and material misstatement pathways.

2. Literature Review

The first is related to analytical evidence and the evolving structure of the audit process around risk. Yoon, Hoogduin and Zhang [5] define Big Data as complementary audit evidence, not as an automatic replacement for existing sources of evidence, and highlight its relevance, reliability and provenance. Technology capability is also cited as one of a number of factors that influence adoption, with organizational incentives, standards, data access, and auditor acceptance influencing implementation, as stated by Alles [6]. Earley [7] observes that the use of audit analytics can increase the population scope and yield better rates of identifying unusual transactions, but there are limitations such as interpretive burden and data-quality issues which may hinder computational gains. Kokina and Davenport [8] describe AI as particularly suitable for automating structured, repetitive audit tasks, whereas more ambiguous procedures involving causality, intent, or materiality continue to require substantial professional judgment. Together, this ensemble suggests that an anomaly's place in the evidence hierarchy in a case is more important than whether it can be detected by an algorithm for effective risk-based auditing. The second cluster explores how audit data analytics is evolving as an engagement technology. Gepp et al. [9] record the growing application of data-intensive methods in audit research and practice, and the failure to routinely use them in practice. Salijeni, Samsonova-Taddei and Turley [10] demonstrate that technology can have as much impact on the nature of the audit process, professional roles and evidential practices as it does on improving current procedures. Engagement evidence is provided by Eilifsen, Kinserdal, Messier and McKee [11] that audit data analytics can be used as an aid to risk assessment and to substantive work, but is limited by data extraction, data quality, documentation, methodological consistency and organizational conditions. The results contradict the notion that the more detailed the information, the more accurate the risk assessment will be. Population-level analysis can expand detection opportunities but also result in larger sets of exceptions for which it is not clear whether the audit is significant. The third group focuses on computational detection of fraud and financial misstatement. Amani and Fadlalla [12] present data-mining applications in the accounting field and document wide applications of the data-mining techniques: classification, clustering, neural and hybrid. While West, Bhattacharya and Islam [13] demonstrate the need for careful consideration of

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 10 October 2026
Page No: 5958 - 5967
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0764

class imbalance, feature relevance, dynamic behaviour and evaluation design for intelligent financial-fraud detection. Lin, Chiu, Huang, and Yen [14] directly compare data-mining techniques with expert judgment in financial-statement-fraud detection, suggesting that consistent financial-statement-fraud detection patterns can be uncovered that are hard to replicate by unaided professional judgment. Hajek and Henriques [15] have also attempted to detect beyond the base numeric accounting variables, by using corporate annual reports combined with machine learning techniques. As relevant as ever, Bao, Ke, Li, Yu and Zhang [16] show that machine learning (ML) can in fact outperform traditional accounting-statistical benchmarks for out-of-sample accounting-fraud detection in the presence of significant class imbalance explicitly.

Table 1 Summary of key findings

Ref	Focus	Key Findings
[5]	Big Data as audit evidence	Expanded data can complement conventional evidence, but evidential value depends on relevance, reliability, provenance, and audit context.
[6]	Adoption of Big Data in auditing	Technical feasibility is insufficient; regulation, incentives, data access, and professional acceptance condition adoption.
[7]	Audit data analytics	Population analysis improves anomaly visibility but increases challenges involving data quality, interpretation, and exception management.
[8]	Artificial intelligence in auditing	Automation is strongest for structured tasks; higher-judgment procedures require human interpretation and contextual evidence.
[9]	Big-data techniques in audit research and practice	Analytical capability has expanded faster than systematic engagement-level implementation and validation.
[10]	Audit technology transformation	Advanced analytics can alter evidential practices, audit organization, and professional roles rather than merely accelerate established tests.
[11]	Audit-data-analytics use on engagements	Practical use supports risk assessment and substantive testing, but extraction, quality, documentation, and implementation barriers remain material.
[14]	Financial-statement-fraud detection	Data-mining classifiers can improve pattern discrimination relative to reliance on expert judgment alone, subject to model-specific trade-offs.
[15]	Annual-report fraud identification	Combining narrative information with financial variables increases the information available for machine-based fraud discrimination.
[16]	Machine learning for accounting-fraud detection	Rare-event machine learning using raw accounting information improves out-of-sample fraud identification relative to conventional benchmarks.

Table 1 shows a structural imbalance in the literature. While the audit studies are strong on evidence, workflow, implementation and judgment, they are relatively weak in providing end-to-end misstatement-classification evidence, whereas the fraud-detection studies are strong on predictive tests and do not appear to provide the same level of attention to the integration of the misstatement classification audit. An AI-based risk-based auditing framework should overcome this gap by using machine-generated anomalies as evidence between the machine and risk, needing to be calibrated, put into context and matched to likely statement-level misstatement.

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 10 October 2026
Page No: 5958 - 5967
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0764

3. Methodology

The methodologies found in literature can be organized on four interrelated layers: audit-risk framing, data and feature building, anomaly or misstatement estimation, and escalation, based on evidence. In contrast to dealing with a chosen model, audit-risk framing starts with the features of the account, assertion, control, and transaction. Then, the analytical layer works on journal entries, ledger balances, financial ratios, corporate disclosures or any other records available to the audit. The current methods include descriptive population analytics, rules-based exception detection, supervised machine learning, data-mining classifiers, neural methods [9] and hybrid text-financial models [12]–[16]. Methodological diversity is useful in that it allows for the detection of different types of abnormality, but comparability is restricted if different target definitions, sampling schemes, class-balancing procedures, validation windows, and performance criteria are used.

Table 2 Method comparison

Ref	Method	Strengths	Limitations
[1]	Big-data analytical audit architecture	Extends audit testing beyond small samples and supports multivariate relationships across large accounting populations.	Conceptual capability does not specify a validated anomaly-to-misstatement decision rule.
[3]	AI-assisted audit formalization	Suitable for repetitive screening and structured exception tasks; supports scalable decision assistance.	Formalization is weakest where materiality, intent, business context, or ambiguous evidence dominates.
[9]	Data-intensive audit analytics	Accommodates large and heterogeneous audit data and permits broad analytical testing.	Empirical engagement validation and standardized performance criteria remain limited.
[12]	Classification, clustering, neural, and hybrid data mining	Provides multiple mechanisms for nonlinear pattern recognition and accounting prediction.	Performance is highly dependent on labels, preprocessing, feature selection, and evaluation design.
[14]	Financial-statement-fraud classifiers	Provides direct fraud-discrimination evidence and explicit comparison with expert judgment.	Fraud labels are rare, historical, and narrower than the full material-misstatement construct.
[15]	Text-financial machine learning	Captures narrative and numeric signals that may reflect manipulation jointly.	Text features can be unstable across reporting styles, industries, and time periods.
[16]	Rare-event ensemble machine learning on accounting data	Addresses severe class imbalance and enables out-of-sample accounting-fraud prediction.	Predictive fraud probability does not by itself identify affected assertions or quantify audit materiality.

Supervised classification provides the most direct link to financial misstatement because the training labels can reflect fraud, statement-level restatement or any other financial statement-level result. One of its main strengths is a direct predictive evaluation against an a priori target, a feature exemplified by financial-fraud research [14]–[16]. The primary limitation is reliance on less than perfect and limited labels, particularly since confirmed fraud is a limited and special subset of material misstatement. Unsupervised anomaly detection has the reverse characteristics: They can be used to rank abnormal transactions without requiring a prior indication of fraud, which makes them appealing for a complete

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 10 October 2026
Page No: 5958 - 5967
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0764

population audit screening, but abnormality is not a misstatement. Audit data analytics falls somewhere between the extremes. Population and multivariate screening can assist in risk assessment and in substantive procedures as a specific exception [7, [11]]; for analytical exceptions, evidence of purpose, authorisation, accounts, cutoff, valuation or disclosure needs to be provided before an audit conclusion can be defended.

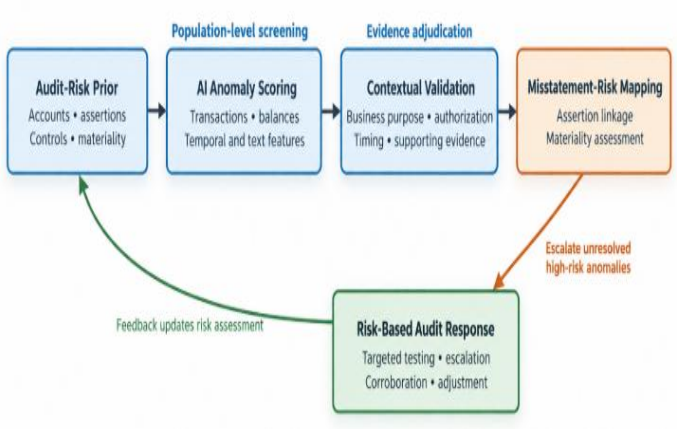

Figure 1 AI-driven risk-based auditing architecture linking anomaly detection with financial misstatement identification

It is therefore necessary to have sequential risk updating, not just one classification event, in a coherent framework. Initial audit-risk priors should guide analytical attention, features of a transaction or account should yield a calibrated score for the anomaly, contextual evidence should guide the interpretation between an explainable exception and an unresolved signal, and unresolved signals should update the risk of misstatement for an assertion. The architecture is shown in figure 1. The framework positions machine-generated assessment of anomalies between the initial audit-risk assessment and the evaluation of the misstatement at the assertion level, instead of as audit conclusions. Risk Prior, transaction-level analytical signals, contextual validation, and assertion-level escalation are part of a closed pathway of evidence (see Figure 1). This is important because it ensures that statistically outlying observations are not automatically considered material financial misstatements and allows model results to be linked to the audit conclusions.

4. Discussion

Studies reported suggest that analytical coverage has been more successful than the integration of audit risk. The population oriented analytics can process far more accounting records than traditional sample-based analytics, with a greater chance of unusual combinations of account, amount, timing, user, or transaction characteristics becoming apparent [1, 7, 11]. This can be very helpful for low frequency variations that are embedded in the larger homogeneous population. However, the quantity of cases that are detected and the quality of the detection are different. Brown-Liburd et al. [2] mention the cognitive and decision issues with large and complex evidence sets, and Eilifsen et al. [11] suggest that engagement deployment is impacted by the availability, extraction, reliability, documentation and usefulness of data. The suggestion is that AI-based auditing should focus on increasing the proportion of decision-relevant flags while controlling the overall exception workload. High-sensitivity systems with many false positives may simply transfer the analytical burden from transaction inspection to exception adjudication without reducing overall audit effort. The second relates to anomaly detection versus financial misstatement identification. While audit analytics will usually detect deviations from expectations, financial-statement-fraud models will attempt to predict a more specific outcome [14]–[16]. This is a technical difference. A statistically unusual transaction may be perfectly ordinary whereas a coordinated statement may be a series of ordinary transactions collated together so as to not meet the simple outlier definition. Therefore, hybrid modelling makes more sense than using a single anomaly statistic. Hajek and Henriques [15] show the importance of using both narrative and financial data and suggest that reporting actions may carry complementary signals that cannot be detected in accounting ratios. Bao et al. [16] also demonstrate that rare-event machine learning can gain predictive information directly from accounting variables and be superior to traditional statistical methods in out-of-sample fraud detection. This evidence is helpful as a tool for machine learning to identify risks but does not stand alone as evidence of misstatement. Fraud labels encompass discovered cases and enforcement

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 10 October 2026
Page No: 5958 - 5967
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0764

results, leading to selective-observation issues; material misstatement covers error and/or other non-fraudulent issues that can have different data signatures. Another tension that has been noted in the literature is interpretability versus performance. Conventional analytic approaches may not be able to state easily the combinations of abnormal values that can lead to fraud risk, but complex models can [12] [13] [16]. But audit evidence must provide more than accurate discrimination. An auditor has to answer a number of questions to determine why an observation is important, what assertion may be impacted, what evidence may support or challenge the signal, and whether the potential impact may be material. This requirement is of importance even in the face of marginally better classification by black-box alternatives, for it gives relatively interpretable feature structures, calibrated probabilities, local explanation techniques and explicit escalation rules practical importance. Thus, four key areas of evaluation need to be considered for the architecture of AI used for auditing purposes: predictive discrimination, calibration, exception workload, and evidential interpretability. The first dimension has been explored extensively in the literature, while the other dimensions have been explored less so.

Table 3 Results comparison

Ref	System / Context	Metric / Basis of Comparison	Outcome
[5]	Big Data used as complementary audit evidence	Evidential relevance and reliability relative to established audit evidence concepts	Expanded data can increase corroborative value, but provenance and reliability determine whether an analytical signal is persuasive.
[7]	Audit-data-analytics applications	Population coverage and exception identification relative to conventional audit procedures	Analytics expands anomaly-search coverage, while interpretation and data-quality problems become central constraints.
[11]	Audit engagements using data analytics	Engagement use, procedural application, and implementation barriers	Analytics supports risk assessment and substantive work, but extraction, quality, documentation, and organizational constraints limit uniform deployment.
[14]	Financial-statement-fraud classification	Data-mining classification compared with expert judgment	Computational models provide useful fraud discrimination and can exceed unaided expert classification for pattern-intensive tasks.
[15]	Corporate annual-report fraud detection	Machine-learning models using financial and narrative information	Combined information sources improve the representation of fraud-related signals beyond single-source financial indicators.
[16]	Accounting-fraud prediction in public firms	Out-of-sample discrimination using rare-event machine learning relative to conventional statistical benchmarks	The ensemble machine-learning approach improves fraud identification and addresses severe class imbalance more effectively than standard benchmark modelling.

As shown in Table 3, the strongest empirical evidence is found when the target is limited, e.g., accounting fraud, while evidence of implementation comes from the audit-engagement studies, with mixed evidence

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 10 October 2026
Page No: 5958 - 5967
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0764

and limited standardization of predictive evaluations. This difference restricts direct comparisons of reported performance. A high-performing fraud classifier and a helpful audit anomaly detector address related but different problems: the first is to predict a statement-level result while the second is to rank records for subsequent examination. Therefore, a risk-based framework should have a minimum of two calibrated outputs: anomaly likelihood at the record level and conditional misstatement risk (after considering audit context).

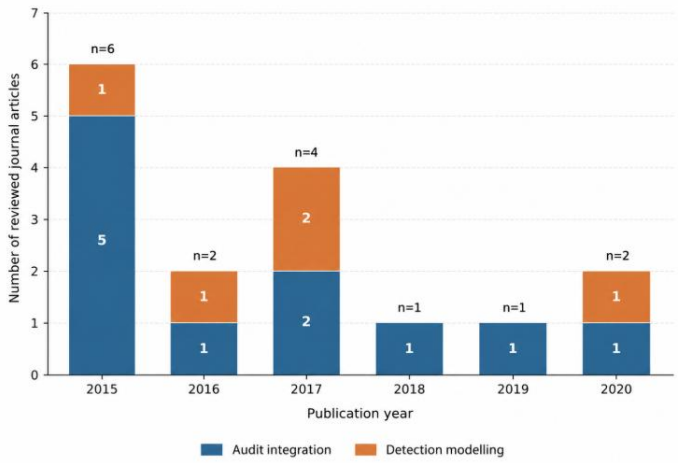

Figure 2 Distribution of reviewed studies by publication year and primary evidence orientation

There is also a disorganized evidence base, as evidenced by temporal distribution. The 16 reviewed articles are divided into audit integration and computational detection categories in Figure 2. This coding is not performance-based as many audit studies lack similar accuracy statistics. The stacked bars separate studies focused primarily on audit integration from studies focused primarily on anomaly, fraud, and/or misstatement detection. As Figure 2 indicates, audit-integration research and computational-detection research evolved side by side and did not coalesce in a quick and tidy empirical stream. This is important because an end-to-end risk-based audit architecture requires evidence to support engagement feasibility, as well as evidence of misstatement discrimination, validated separately from literature.

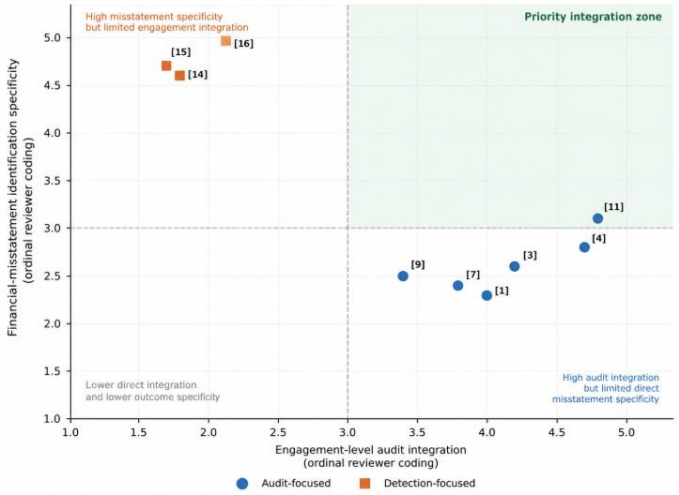

Figure 3 Integrative map of audit-integration evidence and misstatement-identification specificity

An integrative interpretation yields another result: few studies are high on both direct audit integration and direct financial-misstatement specificity. Conceptual audit research is very similar to engagement workflow but may not have transaction-to-misstatement validation [1] [3] [4]. Research on fraud classification is near the misstatement end of the audit-risk spectrum but is often conducted apart from the audit-risk process [14]–[16]. Implementation gap can be filled by engagement research (e.g., Eilifsen et al. [11]), but predictive outcomes are not easily comparable between companies and analytic applications. This structural separation is performed in figure 3 but is not listed as a reported effect estimate as subjective coordinates. Figure 3 depicts each reviewed article plotted in relation to the integration of the audit and the level of engagement and to the degree of specificity in the relationship with identification of anomalies, fraud and/or financial-misstatement. The upper-right, sparse area of Figure 3 indicates research that integrates strong engagement and direct misstatement-identification validation. This gap is important as it is essential that future frameworks are both predictive models that sit within the audit-risk and evidence processes, and engagement analytics that are judged against the financially meaningful outcomes of misstatements.

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 10 October 2026
Page No: 5958 - 5967
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0764

There are a number of technical implications arising from this. First of all, class imbalance is not an easily ignored preprocessing option. Because fraud and material misstatements can occur in transactions that are uncommon in the normal transactions or non-fraud firm-years, a model that predicts the dominant class accurately might lead to a false sense of accuracy when applied to the transactions that are being examined [13] [16]. When analytical outputs result in expensive audit investigation, precision, recall, area-based discrimination metrics, calibration, and workload sensitive measures, are more informative. Second, temporal validation is better than random cross-validation where the intended deployment is likely to misstate the future, since accounting behaviour, standards, controls and economic conditions change over time. Third, different types of audit errors do not have equal costs. The consequences of a missing material misstatement and investigating a benign anomaly are not the same and must be taken into account when determining the threshold. Another caveat is the level of analysis. Audit procedures tend to be conducted at the journal entry, transaction, account or assertion level [11], [14]–[16], while fraud-detection studies typically are conducted at the firm-year level. A firm-year prediction can provide evidence that a company should be studied but not the specific areas of the accounting that are problematic. On the other hand, transaction level anomaly systems can detect anomalies that don't have an overall material impact. It is therefore a key unresolved requirement that multi-level modelling should be implemented. Transaction level anomalies should be aggregated to account level patterns; account patterns should be mapped to assertions like existence, completeness, valuation, rights and obligations, presentation, and cutoff; and the risk of each assertion should then be considered in light of materiality and corroborating evidence. Parts of this architecture are available in the current literature, with sparse empirical support for the overall chain. The evidence reviewed then lent support to the cautious optimism. Using AI to enhance screening breadth and pattern discrimination, and audit analytics to restructure evidence collection and risk prioritisation, can improve the effectiveness of audits [3], [8], [11], [16]. But the quality of the audit is not determined by predictive performance. More evidence would be useful for demonstrating improvements in audit procedure selection, to minimize material misstatements, to keep false positive workload under control, and to ensure consistent traceability and decisions that are robust across entities and reporting periods. Automated anomaly detection should therefore be treated as a risk-ranking and evidence-generating tool rather than as an independent conclusion regarding financial misstatement.

5. Future Directions

More end-to-end validation of AI-based risk-based auditing is needed in the future, instead of more standalone classifiers. A first priority is hierarchical modelling that relates transaction level anomalies to account level risk, financial-statement assertions and material misstatement. While there are frameworks for engagement analytics and fraud-classification research [11] and [16] that provide complementary bases, there is limited empirical evaluation of that hierarchy. Reports on studies should therefore include answers to the questions of whether algorithmic flags result in verified audit adjustments, control deficiencies, restatements, or other outcomes of the engagements, rather than statistical abnormality alone. Validation under realistic, rare event and temporal conditions is the second priority. The research in financial fraud provides evidence for the significance of class imbalance [13, 16] but the research in audit requires metrics that are sensitive to costs, including false-positive investigation effort, severity of missed misstatement, materiality, and audit hours available. Randomly partitioned historical samples would not be as convincing as prospective or rolling window validation. Model calibration should also be an integral part of the requirement, as risk-based auditing requires both correct ordering of observations, as well as interpretable probability or ranking information.

Explainability and evidential traceability are a third priority. AI outputs should highlight influential attributes of transactions, associated accounts, temporal trends, and likely assertion effects, to enable subsequent processes to be designed directly. AI formalization and audit analytics studies suggest that the benefits of automation are found when

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 10 October 2026
Page No: 5958 - 5967
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0764

computational outputs are still used as part of defensible professional processes [3] [4]. Explainability should thus be evaluated not as a mere software property, but also as an audit-performance element. Finally, investigating adaptive monitoring in evolving enterprise systems is important. If the ERP is migrated, controls are redesigned, an entity is acquired, new regulations are introduced, or transaction patterns change, the models may become invalid. Empirical audit studies should explicitly involve drift detection, control of models, reproducible data pipelines, independent validation, and secure handling of client data. Such instructions would take risk-based auditing based on AI from anomaly engines to an AI-enabled system of continuously governed financial-statement risk.

Conclusion

AI-based risk-based auditing is the result of several parallel, but not fully connected, streams of research. These audit-analytics and artificial-intelligence research efforts help to validate the benefits of increased transaction coverage, population-based testing and automated exception screening, and formalizing and scaling structured audit activities. Research on financial misstatement and fraud detection shows that multivariate reporting patterns can be identified by machine learning models, which are not necessarily detected by classical statistical or judgment models. The principal scholarly hurdle is no longer establishing that algorithms can identify atypical accounting activity. The more significant problem is developing a defensible chain of reasoning from abnormality to audit risk, assertion-level evidence, materiality, and financial misstatement. The current research is limited by labels that are rare, inconsistent, followed by a burden of false positive results, inaccessible models, dynamic data distributions, and the lack of appropriate units of analysis for the transaction-level audit tests and the firm-level fraud prediction. A good framework, therefore, should integrate risk priors, anomaly scoring based on calibration, validation of the context, multi-level aggregation, mapping assertions, and escalation based on evidence. The most defensible use of AI is as a tool to prioritize and enhance audit evidence not to supplant professional evaluation. Progress will be dependent upon empirical designs that will test all audit pathways and assess predictive performance along with interpretability, workload, calibration, and verified misstatement results.

References

[1]. Cao, M., Chychyla, R., & Stewart, T. (2015). Big Data analytics in financial statement audits. *Accounting Horizons, 29*(2), 423–429.

[2]. Brown-Liburd, H., Issa, H., & Lombardi, D. (2015). Behavioral implications of Big Data's impact on audit judgment and decision making and future research directions. *Accounting Horizons, 29*(2), 451–468.

[3]. Issa, H., Sun, T., & Vasarhelyi, M. A. (2016). Research ideas for artificial intelligence in auditing: The formalization of audit and workforce supplementation. *Journal of Emerging Technologies in Accounting, 13*(2), 1–20.

[4]. Appelbaum, D., Kogan, A., & Vasarhelyi, M. A. (2017). Big Data and analytics in the modern audit engagement: Research needs. *Auditing: A Journal of Practice & Theory, 36*(4), 1–27.

[5]. Yoon, K., Hoogduin, L., & Zhang, L. (2015). Big Data as complementary audit evidence. *Accounting Horizons, 29*(2), 431–438.

[6]. Alles, M. G. (2015). Drivers of the use and facilitators and obstacles of the evolution of Big Data by the audit profession. *Accounting Horizons, 29*(2), 439–449.

[7]. Earley, C. E. (2015). Data analytics in auditing: Opportunities and challenges. *Business Horizons, 58*(5), 493–500.

[8]. 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.

[9]. Gepp, A., Linnenluecke, M. K., O'Neill, T. J., & Smith, T. (2018). Big data techniques in auditing research and practice: Current trends and future opportunities. *Journal of Accounting Literature, 40*, 102–115.

[10]. Salijeni, G., Samsonova-Taddei, A., & Turley, S. (2019). Big Data and changes in audit technology: Contemplating a research agenda. *Accounting and Business Research, 49*(1), 95–119.

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 10 October 2026
Page No: 5958 - 5967
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0764

[11]. Eilifsen, A., Kinserdal, F., Messier, W. F., Jr., & McKee, T. E. (2020). An exploratory study into the use of audit data analytics on audit engagements. Accounting Horizons, 34(4), 75–103.
[12]. Amani, F. A., & Fadlalla, A. M. (2017). Data mining applications in accounting: A review of the literature and organizing framework. International Journal of Accounting Information Systems, 24, 32–58.
[13]. West, J., Bhattacharya, M., & Islam, R. (2016). Intelligent financial fraud detection: A comprehensive review. Computers & Security, 57, 47–66.
[14]. Lin, C.-C., Chiu, A.-A., Huang, S. Y., & Yen, D. C. (2015). Detecting the financial statement fraud: The analysis of the differences between data mining techniques and experts' judgments. Knowledge-Based Systems, 89, 459–470.
[15]. Hajek, P., & Henriques, R. (2017). Mining corporate annual reports for intelligent detection of financial statement fraud: A comparative study of machine learning methods. Knowledge-Based Systems, 128, 139–152.
[16]. Bao, Y., Ke, B., Li, B., Yu, Y. J., & Zhang, J. (2020). Detecting accounting fraud in publicly traded U.S. firms using a machine learning approach. Journal of Accounting Research, 58(1), 199–235.