The role of random forest in internal audit to enhance financial reporting accuracy

Eid M. Alotaibi[a*], Ashraf Khallaf[b] and Kimberley Gleason[c]

[a]*Assistant Professor, Department of Accounting, American University of Sharjah, United Arab Emirates*
[b]*Professor, Department of Accounting, American University of Sharjah, United Arab Emirates*
[c]*Professor, Department of Finance, American University of Sharjah, United Arab Emirates*

CHRONICLE	ABSTRACT
	Internal audit is a bulwark ensuring the integrity of financial statements, a linchpin for stakeholder trust and informed corporate decision-making. With the proliferation of complex financial transactions, audit teams face mounting challenges in deciphering voluminous transactional data to safeguard financial reporting quality. Machine learning has the potential to identify signifiers of financial reporting quality. Within the Design Science Methodology framework, we apply the Random Forest Classifier technique to metrics such as the error rate to enhance financial reporting. We find that the Random Forest Classifier identifies that certain parameters are critical to error detection, which enhance account receivable accuracy, lower receivable account control risk. This research advances the argument that technologically-enhanced internal audit procedures can play a pivotal role in ensuring that financial reporting mirrors the economic reality of the company.

1. Introduction

The internal audit function is the third line of defense in shoring up an entity's control environment. It is a foundational pillar safeguarding the integrity of financial statements, ensuring stakeholders are presented with an unambiguous snapshot of an organization's fiscal health. These statements guide corporate decision-making and anchor investor trust, and the importance of the accuracy of financial statements cannot be overstated. Accounting Information Systems (AIS) have ushered in a new era of financial data processing, merging the disciplines of accounting and information technology (Romney et al., 2012). The vastness and intricacy of transactions, the increasing sophistication of reporting systems, and the complexity of reporting standards, demand cutting-edge tools for interpretation and scrutiny of financial and accounting data (Stoel et al., 2012). The primary mandate of internal audit teams revolves around identifying, rectifying, and preventing financial transaction errors. With the complexity of today's financial ecosystems, traditional auditing tools and methodologies may yield type 1 and 2 errors in the detection of discrepancies. Minor discrepancies, when viewed in isolation, might appear inconsequential. However, such errors (particularly, type 2 error: incorrect acceptance) can cumulatively result in material misstatements, obfuscating the information presented by financial statements regarding the financial stability of an organization. The richness of data generated by AIS has made manual oversight not only tedious but almost impracticable. Accordingly, new technologies, such as Machine Learning (ML) can be applied for error detection, used to minimize the oversight of transactions that may otherwise yield errors that could diminish stakeholder confidence or trigger regulatory consequences, as observed in past corporate debacles where overlooked clusters of errors contributed to significant financial misrepresentations (Healy & Palepu, 2003).

In this paper, we use the Design Science Methodology (DSM) framework to address challenges faced by internal audit teams as they navigate labyrinthine AIS environments in their pursuit to ensure accuracy in financial reporting. To do so, we describe and apply the Random Forest Classification (RFC) machine learning technique and demonstrate its application to the improvement of the precision and efficiency of audits. The RFC, with its robust predictive capability, has been acknowledged for its capabilities for improving audit efficiency (Breiman 2001). By embedding these computational strengths within the DSM framework, we seek to merge technological advancements with pragmatic audit imperatives. We find that the RFC, when applied to data containing information regarding accounts receivables and customer characteristics can identify clusters of errors which in turn yields a reduction in error and an improved Financial Reporting Quality Score (FRQS).

We contribute to the literature in three key ways. First, this paper is the first to incorporate the RFC into the DSM framework. Second, we are the first to apply the RFC to error detection and demonstrate the corresponding impact on financial reporting quality. We show that this approach can identify not just individual transaction errors but to identify and rectify clusters transaction errors. Third, as financial paradigms evolve and data repositories swell, it is critical for the audit framework to evolve; accordingly, we offer new suggestions and recommendations for managers.

The paper proceeds as follows. In section 2, we review the relevant research on the integration of ML into the DSM for the purpose of improving financial reporting quality. Section 3 describes the techniques and strategies employed, establishing the robustness of our research process. In section 4, we subsequently discuss the implications of our findings, connecting the dots between real-world applications and our results. The final section concludes the paper, offering suggestions for the audit function as well as future research.

2. Literature Review

Given the pace at which technological advancements are being made in AIS, it is important to consider how emerging technologies and methodologies fit in with traditional auditing practices. There are three relevant areas of research required to explore the integration of ML into audit and financial reporting quality: the internal auditing of financial transactions, machine learning's transformative potential in auditing, and DSM as aligned with objectives of AIS research.

2.1 Internal Auditing and Financial Transactions Audit

Internal auditing, an integral component of the current data driven business environment, has undergone considerable evolution over time in terms of both scope and objectives, in large part resulting from the increasing difficulties businesses face in upholding financial transparency while meeting regulatory standards. Internal auditing was once seen solely as an audit on financial records; today it serves both compliance and advisory roles to protect organizational financial integrity by verifying that transaction records are accurate, transparent, and compliant with regulations. Roussy et al. (2020) argues for the critical compliance and advisory roles occupied by internal audits in terms of guaranteeing that records remain accurate while meeting regulatory standards.

Internal auditors face two primary challenges when conducting internal audits: first is the complexity of current business processes and transactions that they generate, which; this makes transaction validation an arduous task; second is an exponential rise in the volume of available data. AIS often generates huge volumes of transaction records, making auditors' tasks almost like locating a needle in a haystack.

Internal auditors utilize substantive testing to detect transaction errors in organizations (Groomer and Murthy 2018). To do so, as part of their quantitative methods, they utilize a sampling strategy, reviewing only selected transactions instead of conducting an exhaustive examination. However, as organizations more complex, with millions of transactions taking place daily, however, this approach may be inadequate; Barr-Pulliam et al. (2022) notes that sampling can miss substantial errors when applied over vast databases and Beven and Binley (1992) further indicate that even seemingly minor individual errors propagate over time and can become material, affecting financial reporting quality and decision making. One of the main objectives of the Committee of Sponsoring Organizations (COSO) internal control framework is to produce reliable financial reports. By accessing and utilizing the entire population of transactions, rather than only a sample of transactions, ML can offer additional tools for more sophisticated error detection before the problem spirals, which in turn decreases sample risk and improves audit quality. .

2.2 Machine Learning in Auditing

In the landscape of technological interventions available to auditors, ML, a vibrant offshoot of artificial intelligence, stands prominent, and has been heralded offering significant transformations in the auditing arena. Brown-Liburd et al. (2015) emphasized its potential, stating that machine learning can radically revamp the efficiency, accuracy, and depth of auditing processes by supporting auditors in deciphering anomalies, detecting patterns, and recognizing potential transaction errors.

Vasarhelyi et al. (2015) describe the benefits of ML applications to audit in terms of its ability to navigate vast data sets, unveiling patterns and inconsistencies that might evade traditional analysis.

The merits of integrating ML into auditing are manifold. Kokina and Davenport (2017) argue that an advantage of ML is its unparalleled prowess in large-scale data processing. Traditional auditing tools, crafted for modest datasets, are overwhelmed when they grapple with the massive data sets produced by contemporary AIS. However, sophisticated ML models, such as the Random Forest algorithm, Neural Networks, and Deep Learning, thrive in such environments. Cho et al. (2020) noted that these algorithms can not only handle enormous volumes of data but can also deduce insights that human auditors might miss.

ML is a paradigm shift in audit practice that is not merely confined to the refinement and augmentation of audit tools; it signals a watershed moment in the methodological trajectory of AIS research (Geerts, 2011). Bardelli et al. (2020) posit that the infusion of machine learning algorithms into AIS can dramatically elevate the precision, efficiency, and reliability of audit procedures. Furthermore, this technological integration has the potential to transcend traditional audit boundaries, paving the way for real-time auditing and continuous assurance mechanisms (Chan et al., 2018). In essence, the AIS discipline stands on the cusp of a monumental transition, underscored by the confluence of advanced computational techniques and foundational accounting principles (Vasarhelyi et al., 1991).

2.3 Random Forest Classifier

The Random Forest classifier (RFC) is a machine learning technique recognized in the literature for its capabilities in gauging classifier performance (Costa et al., 2022). Impressively, the model demonstrated an acute precision in its classifications. Breiman (2001) describes the use of the RFC in error detection. An and Suh (2020) find that the modified RFC improves identification of financial statement errors and fraud. In the context of internal auditing, the application of RFC has proven to be invaluable. Notably, An and Suh (2020) have shown that the modified RFC enhances the identification of financial statement errors and fraud. This underscores the potential of RFCs to contribute significantly to the field of internal auditing by improving the accuracy and efficiency of anomaly detection in financial data. The use of RFC in internal auditing can lead to more efficient, accurate, and proactive auditing processes. By harnessing the power of machine learning, auditors can better identify risks (by analyzing historical data, auditors can predict which transactions are more likely to be associated with errors, irregularities, or fraud), detect anomalies or irregularities (by identifying unusual patterns or outliers), and enhance the overall effectiveness of their audit efforts. However, it's crucial for auditors to understand the technology, properly implement it, and interpret its results to reap the full benefits of utilizing RFC.

3. Methodology

We employ Design Science Methodology as a framework to navigate the intricacies of AIS and integrate accounting principles with the transformative capabilities of contemporary technology. The synthesis of DSM with AIS research offers a new dimension to internal audit research. Hevner et al. (2004) and Gregor and Hevner (2013) describe DSM as a methodology rooted in the creation and systematic evaluation of artifacts. These artifacts are constructed to address and potentially solve specific organizational challenges.

Table 1
The Key Steps of the DSM in AIS

Step	Description	Relevance in AIS
1. Problem Identification	Discerning and articulating a real-world issue necessitating innovation.	Unearthing challenges, inaccuracies, or inefficiencies in the accounting procedures steered by prevailing systems.
2. Define Objectives for a Solution	Charting out desired outcomes and intentions for the envisioned solution.	Determining what an AIS tool or intervention should ideally achieve, be it in precision, operational efficiency, or user compatibility.
3. Framework Design and Development.	Conceptualizing a preliminary solution and progressively refining it.	Crafting an AIS solution and developing a framework.
4. Demonstration	Validating the solution's operability in actual or quasi-real scenarios.	Deploying the AIS solution in genuine accounting settings to demonstrate its utility and prowess.
5. Evaluation	Rigorously gauging the efficacy of the solution against its predefined objectives.	Analyzing the AIS tool's alignment with initial aspirations, using error detection.
6. Discussion	Disseminating the acquired knowledge—both the challenge and its remedy—to pertinent audiences.	Sharing research insights and practical takeaways with the wider AIS community, ensuring shared growth for both practitioners and academics.

Adapted from (Hevner et al. 2004; Gregor and Hevner 2013)

In essence, DSM straddles the theoretical and the applied, serving as a bridge. David et al. (2002) and Geerts (2011) opine that from a theoretical perspective, DSM is compatible with AIS research as AIS traditionally aims to design, implement, and critically assess information systems tailored to cater to accounting and auditing needs, and DSM, with its structured framework, reinforces this endeavor by ensuring that the systems generated are both innovative and practically effective.

Reconciling machine learning techniques within this DSM framework, as our research proposes, marks a significant leap in that indicates an evolution, not just in the tools used for auditing, but also in the methodological underpinnings of AIS research. The DSM in AIS offers a systematic approach that integrates theoretical investigation with pragmatic application. By doing so, the DSM ensures that research outcomes resonate with both academic and practical challenges in the realm of AIS. Table 1 describes the operationalization of the DSM, shedding light on its relevance to AIS. The DSM process equips researchers with a holistic perspective, starting from the germination of a problem to the dissemination of the solution. In the AIS landscape, the DSM's structured approach is critical because addressing intricate issues like financial transaction errors requires a clear roadmap. The DSM provides this blueprint, guiding researchers through methodical exploration while ensuring that the resultant findings and tools hold tangible value for the AIS practitioner community.

3.1 Problem Identification and Definition

The objective of this study is to illustrate the utility of the Random Forest Algorithm in financial statement error detection, a persistent challenge for internal auditors. While individual transaction errors might appear inconsequential, a collective examination often reveals patterns that can translate into significant financial discrepancies (Studer et al., 2021). This challenge becomes increasingly daunting with the escalating complexity and volume of financial transactions in contemporary business environments. One major factor exacerbating this problem is the present transition of businesses from traditional brick-and-mortar operations to intricate digital ecosystems. This digital transformation has expanded the scope and scale of financial transactions and introduced new dimensions of variability and unpredictability in transactional data (Cohen et al., 2010). Another facet is the limitations inherent in traditional auditing methodologies (Byrnes et al., 2018). As previously highlighted, substantive testing and sampling, though adequate for smaller datasets, often miss critical errors in voluminous datasets (Groomer & Murthy, 2018). These factors merge to create a scenario where detecting financial transaction errors becomes akin to searching for a needle in a haystack. However, merely identifying the magnitude of the problem is insufficient. To be of tangible value, the problem must be defined in a manner that guides the subsequent phases of research. The problem definition adopted in this study, therefore, is twofold:

1. To explore the limitations of current auditing methodologies in accurately detecting financial transaction errors in technology-driven business environments.
2. To investigate the Random Forest algorithm as a potential technological and methodological intervention that can enhance the accuracy and efficiency of financial transaction error detection.

In the realm of accounting information system research, exploring the application of Random Forest algorithms offers a new approach to the examination of financial data accuracy in the age of the enterprise resource planning system (ERP). Rather than relying on the tried-and-true, but limited, sampling methods, incorporating this machine learning technique allows us to examine the full spectrum of a company's transactions, shifting the detection of financial errors from a rearview mirror to a high-definition, panoramic scope of financial accuracy and risk. The Random Forest can redefine the detection of irregularities and bolster the trustworthiness of financial report, advancing the understanding of the subtleties involved in modern financial systems amid a technologically advanced landscape.

The significance of a well-articulated problem sets the trajectory for the research, ensuring that every subsequent step – from data collection to analysis and interpretation – is aligned with the research's core objectives (Turkay et al., 2017). Moreover, in the realm of AIS, where the confluence of accounting practices, organizational needs, and technological innovations is at play, a clear problem definition becomes the anchor, preventing the research from veering into tangential directions and ensuring that it remains relevant and actionable for both practitioners and researchers.

The problem identification and definition phase were both a reflective and systematic exercise, carefully drawing from existing literature, recognizing the gaps and challenges in current auditing practices, and subsequently sculpting a well-defined research trajectory that promises actionable insights and novel contributions to AIS.

3.2 Define Objectives for a Solution

The cornerstone of any rigorous research approach, particularly in the realm of AIS, lies in the delineation of clear and achievable objectives for the envisaged solution. These objectives serve as the guiding light, directing researchers toward creating solutions that are not only innovative but also address the pertinent challenges inherent in the domain (Gregor & Hevner 2013). In the context of AIS, financial transaction errors are increasingly becoming focal points of concern for businesses globally. As a result, identifying and rectifying these errors has become paramount, necessitating solutions that can efficiently detect and mitigate such inaccuracies (Alles, 2015). It is within this backdrop that we endeavor to define our solution objectives, ensuring alignment with the real-world challenges faced by businesses today.

Objective 1: Enhancing Error Detection Accuracy. A paramount objective is to enhance the precision with which financial transaction errors are identified. Traditional methods often grapple with large data volumes, complex transaction structures, and evolving accounting standards, leading to potential oversights. With the integration of advanced technologies, especially machine learning algorithms such as Random Forest, the aim is to significantly reduce false positives and negatives, ensuring that genuine errors are pinpointed with greater accuracy (Cho et al., 2020).

Objective 2: Improving Efficiency of the Audit Process. Time is a valuable resource in the audit realm. While accuracy remains pivotal, enhancing the speed at which transaction errors are identified and rectified can lead to considerable cost savings and reduced business risks. Leveraging machine learning can aid in streamlining this process, analyzing vast datasets swiftly and providing timely feedback to auditors (Sun, 2019).

Objective 3: Facilitating User-friendly Interface and Feedback Mechanisms. A robust AIS solution should not only be technically adept but also user-friendly. Auditors, many of whom might not have in-depth technical expertise, should find the system intuitive. The objective is to design a solution that is easily navigable, offers clear feedback on identified errors, and provides guidelines on rectification steps, bridging the gap between advanced technology and its practical application (Geerts, 2011).

Objective 4: Ensuring Scalability and Adaptability. Given the dynamic nature of the financial landscape, marked by evolving regulations, business practices, and technologies, an ideal solution should be both scalable and adaptable. As businesses grow, transaction volumes and complexities may also increase. The AIS tool should be capable of seamlessly scaling up while adapting to new transaction types or changing accounting standards (Alles et al., 2021).

Objective 5: Fostering Continuous Learning and Adaptation. Machine learning models thrive on data. An essential objective for the solution is to integrate mechanisms that allow continuous learning. As the system encounters newer transaction patterns or novel errors, it should adapt, refining its algorithms to stay abreast with the changing landscape, ensuring sustained accuracy over time (Holzinger et al., 2018).

Each objective outlined above is essential in crafting an all-inclusive solution tailored for the AIS domain. It is critical to remember that these objectives don't operate independently of each other; rather, they work hand-in-hand to form an inclusive tool. For instance, improving detection accuracy (Objective 1) without prioritizing user-friendliness of interface (Objective 3) might present practical implementation challenges, while prioritizing efficiency without continuous learning (Objective 2) may render its relevance obsolete over time.

As we navigate the complex tapestry of AIS research, defining clear objectives for any envisaged solution provides a crucial guide. By setting realistic yet academically rigorous goals that address real business issues (Romney et al. 2012), this ensures the tool or system developed will not only be innovative but will also address significant business needs.

3.3 Framework Design and Development

Central to the DSM in AIS research is the development of a framework tailored to address specific problems identified in the prior stage. By leveraging a structured framework, AIS researchers can better align their strategies with overarching objectives and create tools or systems that adeptly address real-world challenges. A well-designed framework not only provides a blueprint for the entire research process but also acts as a roadmap for implementing the envisioned solution (Gregor & Hevner 2013; Hevner et al., 2004). For this study, the framework is anchored around detecting financial transaction errors, which is becoming an increasingly significant concern within AIS. Given the rapidly transforming financial landscape and the voluminous transaction data generated daily, there is a dire need for an advanced, agile, and adaptive solution.

With this context, we introduced the Financial Transaction Error Detection Framework (FTEDF), conceptualized as a multi-layered, machine learning-driven system tailored for precise and efficient error detection in financial transactions. Fig. 1 shows the FTEDF stages and there are five layers. These layers are in order based on their functioning and they are as follows:

Layer 1 - Data Input: This is the entry point of the framework where raw financial transaction data, from multiple sources, are ingested. Considering the heterogeneity and vastness of financial data, integrating data from various systems and formats is pivotal (Romney et al., 2012). In our study, we pulled receivables and a comprehensive list of customers with designated credit limits.

Layer 2 - Pre-processing: Before analysis, data needs to be cleaned, standardized, and transformed to ensure that it's in a suitable format for machine learning algorithms. This stage also involves the handling of missing data, outliers, and potential discrepancies that could skew results (Alles, 2015).
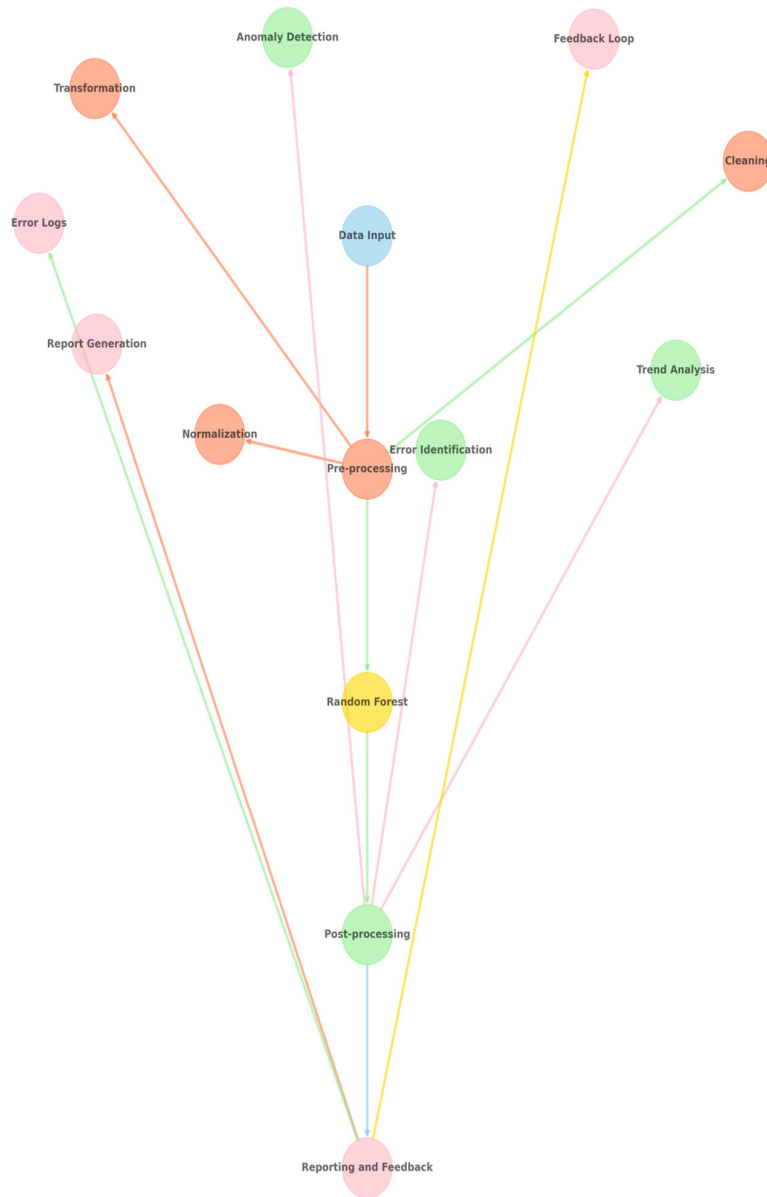
Fig. 1. Financial Transaction Error Detection Framework (FTEDF)

Layer 3 - Machine Learning Algorithms: This layer integrates advanced algorithms, with a specific emphasis on the Random Forest Classifier, to detect potential transaction errors. Random Forest's ensemble learning technique, which employs multiple decision trees and aggregates their results, has been shown to provide high accuracy in diverse datasets, making it a suitable choice for this application (Cho et al., 2020).

Layer 4 - Post-processing: Once errors are identified, they are collated, categorized based on severity, and presented in a format conducive to further analysis. It's imperative that false positives are minimized at this stage to ensure that the audit process remains efficient (Sun, 2019).

Layer 5 - Reporting and Feedback: The final layer involves reporting detected errors to auditors, accompanied by potential rectification steps and insights on error patterns. Furthermore, feedback loops are integrated, enabling continuous refinement of the model based on auditor feedback and evolving transaction patterns (Geerts, 2011).

The FTEDF, as conceptualized, encapsulates a comprehensive approach towards error detection in financial transactions. By leveraging cutting-edge machine learning techniques and ensuring meticulous data handling at each layer, the framework is poised to enhance both the precision and efficiency of the audit process. Moreover, by incorporating feedback mechanisms,

the FTEDF exhibits adaptability, ensuring it remains relevant even as financial practices evolve. Progressive refinement is inherent to the FTEDF's design ethos. As the framework is piloted in real-world scenarios, iterative feedback informs its evolution, optimizing both its technical infrastructure and its alignment with auditor requirements. By embracing a cyclical development process, the FTEDF ensures sustained relevance and efficacy in the dynamic world of AIS (Alles et al., 2021). The proposed framework represents a concerted effort to elevate the accuracy, efficiency, and adaptability of error detection in financial transactions. As AIS research continues to intersect with advanced technologies, frameworks like the FTEDF offer a promising glimpse into the future of financial auditing, characterized by heightened precision and agility (Cho et al., 2020).

3.4 Demonstration of the FTEDF

The DSM in the realm of AIS provides a scaffold, reinforcing the importance of empirical demonstrations in a real-world context (Gregor & Hevner 2013). With the confluence of burgeoning technological advancements and conventional accounting principles, a pressing need arises for frameworks that seamlessly integrate the two, driving both accuracy and efficiency. It is within this backdrop that the FTEDF emerges, asserting its potential as a transformative tool in the intricate landscape of financial error detection. This section aims to meticulously elucidate the application and utility of the FTEDF, employing two realistic datasets to shine a spotlight on its prowess. The FTEDF is designed to holistically evaluate the quality of financial reporting by scrutinizing the underlying data. Both datasets were ingested into the FTEDF. Given the constraints of this demonstration, we have fine-tuned the FTEDF model to capture and analyze metrics specific to account receivables and their corresponding customer credit limits. These metrics are:

3.4.1 Datasets & Data Input

Any robust financial analysis is underpinned by the authenticity and precision of its foundational data. The initial phase in the FTEDF involves meticulous data cleaning and organization. Given the complexities and nuances associated with Excel formatted data, this stage ensures that anomalies, potential data integrity issues, and formatting discrepancies are swiftly flagged and addressed, setting the stage for subsequent phases (Canhoto & Clear, 2020). For this demonstration, we limited our focus to two datasets: (1) a single account – (Account Receivables), and (2) the associated controls present in the customer list, which includes their respective credit limits. The datasets used in this research are synthetic, designed to closely resemble real-world financial transactions without compromising on confidentiality or authenticity. This approach ensures ethical considerations and data privacy remains paramount, while still providing a robust foundation for meaningful evaluation. The Customer Master List, a sample of which is shown in Table 2, is a repository of information for a cross-section of 87 customers. It includes data regarding customer fiscal standing and credentials, which are essential for subsequent stages of analysis.

Table 2
Sample from Customer Master Listing

ACC_ID	CUST_NAME	ADDRESS	CREDIT_LMT
A001	Dan XXXXXX	Company Name, Street Number & Name, City, State, Zip code, Country	20,000

The Account Receivables records, illustrated in Table 3, encompasses 253 records and is used to demonstrate the error detection capabilities of the RFC.

Table 3
Sample from Account Receivables Records

ACCOUNT_NO	INVOICE_NO	GROSS_AMT	GST	PST	PAID_FLAG	DATE_DATE	DATE_TIME	CUST_REF
C020	46000	2345.54	153.45	173.74	-	11/4/14	13:20:10	A5574

3.4.2 Pre-processing

Data is often messy, and this messiness can obfuscate genuine insights. The pre-processing phase emphasizes rigorous data normalization, ensuring consistency and uniformity across records. It is a harmonizing step, transforming disparate data snippets into a cohesive whole. In this step, any potential anomalies or data integrity issues are flagged and rigorous data normalization is pursued. This is a pivotal step, ensuring that subsequent analyses are not plagued by data inconsistencies. Fig. 2 illustrates the cleansing phase outcomes.

```
Cleaned Customer Master Data:
   ACC_ID    CUST_NAME                                                ADDRESS  \
0   A001     DAN ACKROYD   Audenshaw                          125 New Stree...
1   A123     MIKE ATSIL    The Vetinary House                 123 Dog Row  ...
2   A128     IVAN AKER     The Old House                      Ottawa       ...
3   B001     KIM BASINGER  Mesh House                         Fish Street  ...
4   B002     RICHARD BURTON Eagle Castle                      Leafy Lane   ...

   CREDIT_LMT
0      20000
1      20000
2      10000
3      12000
4       9000

Cleaned Receivables Data:
   ACCOUNT_NO  INVOICE_NO TYPE  GROSS_AMT     GST     PST  PAID_FLAG  \
0     C020       46000     I     2345.54   153.45  173.74       NaN
1     C020       46006     I     3039.84   198.87  225.17       NaN
2     M123       46010     I       79.91     5.23    5.92       NaN
3     T004       46010     I       39.94     2.61    2.96       NaN
4     T003       46013     I     3543.54   231.82  262.48       NaN

   DATE_DATE  DATE_TIME  CUST_REF COMMENT
0  2014-11-04  00:00:00     A       5574    NaN
1  2014-12-05  00:00:00     X       9574    NaN
2  2014-12-22  00:00:00     V        587    NaN
3  2015-01-22  00:00:00     Q       8882    NaN
4  2015-01-22  00:00:00     D       8050    NaN
```

Fig. 2. Cleansing Phase Outcomes

3.4.3 Data Integration & Error Detection

This step is a primary indicator of the accuracy of the transaction entries, and it has two phases. In the first phase, each client is grouped with their respective transactions. In the second phase, we merge the A/R dataset with the customer master listing dataset and identify errors of unmatched accounts, as shown in Fig. 3. We then calculate the error rate by dividing the number of unmatched accounts by the total client accounts.
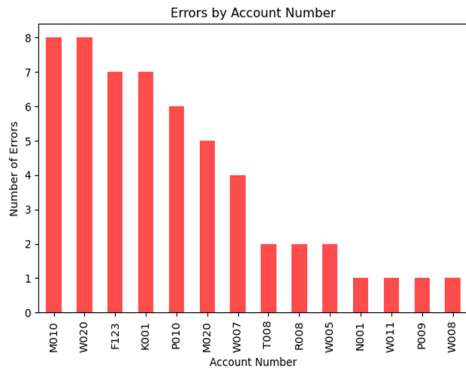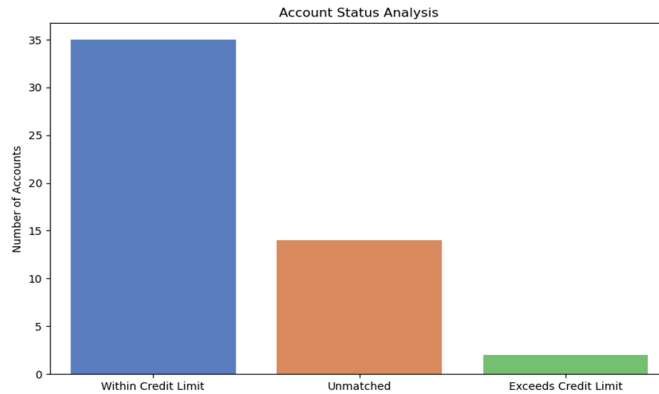

Fig. 3. Errors by Client's Records

Fig. 4. Exceed Credit Limit Records

3.4.4 Exceeding Credit Limit Contribution

We next examine accounts that frequently surpass their credit limit. Surpassing credit limits is a flag of weak internal controls in two regards: poor internal control mechanisms and aggressive revenue recognition strategies. By comparing the gross amount in the account receivables against the credit limit in the customer file, we can identify and count instances where the credit limit is exceeded, calculated as the ratio of the number of exceeded credit limit records to total client accounts. The outcome of this analysis is shown in Fig. 4.

3.4.5 Invoice Volume Analysis

The volume of transactions can be indicative of the business's operational complexity. By tallying the total number of invoices or transactions, we get a sense of the volume. A higher volume, maintained with a consistent error rate, is indicative of robust financial processes. The invoice volume score was calculated by dividing the number of unmatched accounts and the exceeded credit limit by the total number of invoices.

3.4.6 Average Days Outstanding Analysis

This metric gauges how long, on average, it takes for a company to collect payment after a sale has been made. A higher average day's outstanding can be indicative of potential financial distress, ineffective credit and collection procedures, or revenue recognition fraud. By calculating the difference between the invoice date and the payment date for each transaction, then taking an average, we determined the average days outstanding as shown in Fig. 5.
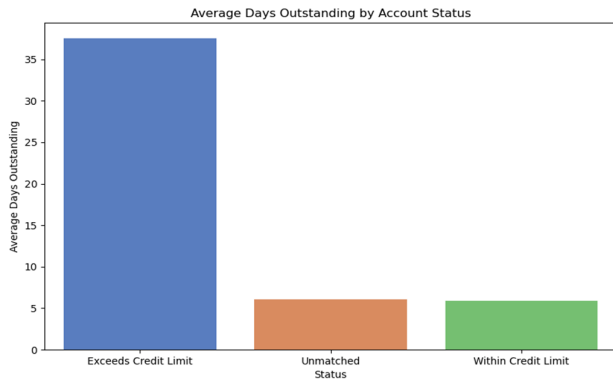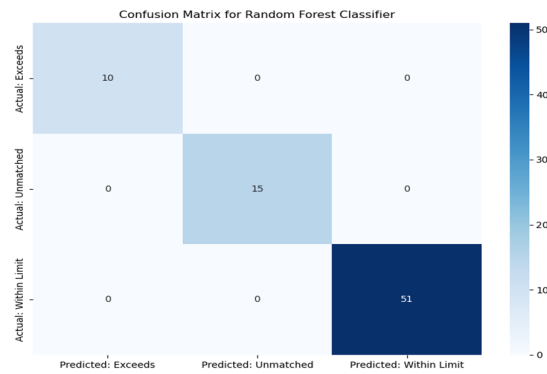
Fig. 5. Average Days Outstanding Analysis

Fig. 6. Classification Matrix Analysis

3.4.7 Random Forest Classifier

Utilizing Random Forest classifier approach, we have developed the Financial Reporting Quality Score (FRQS) as a core component of the Random Forest classifier stage. After determining previous metrics, we amalgamate them into a singular score using the following model:

Financial Reporting Quality Score (FRQS)= w1(Error Rate Score)+w2(Exceeding Credit Limit Score)+w3(Invoice Volume Score)+w4(Average Days Outstanding Score), Where w1, w2, w3, and w4 are the respective weights for each metric, emphasizing their significance in the overall financial reporting quality of the company. The allocation of these weights can be adjusted based on domain-specific knowledge, expert inputs, or iterative testing to ensure that the FRQS is indicative of the actual financial reporting quality.

The incorporation of the FRQS components as features in the Random Forest classifier framework yielded significant insights into the dynamics of financial reporting quality, namely, feature importance and classification accuracy. We demonstrate that the integration of the FRQS with a Random Forest Classifier demonstrates its power in terms of categorizing accounts that exceeded credit limits, unmatched accounts, and those adhering to credit thresholds. Figure 6 shows that, for accounts that exceeded credit limits, we observed an impeccable classification precision, with 10 accounts accurately identified. The unmatched accounts witnessed a similar precision, with 15 correctly categorized. Most commendable, accounts operating within the credit thresholds stood at a resounding count of 51, all correctly pinpointed by the classifier. These findings align with Costa et al. (2022) on the classification precision of the RFC. Fig. 6 visually elucidates these conclusions. Our findings bridge the intricacies of the post-processing layer of the FTEDF and the broader evaluation of the FTEDF and paint a profound portrait of the hierarchy of determinants in the financial transaction landscape. Following our delve into the classification matrix, we next turn our analytical lens towards the feature importance metric, an instrumental tool within the Random Forest framework renowned for its discerning capacity in isolating pivotal features (Breiman, 2001).
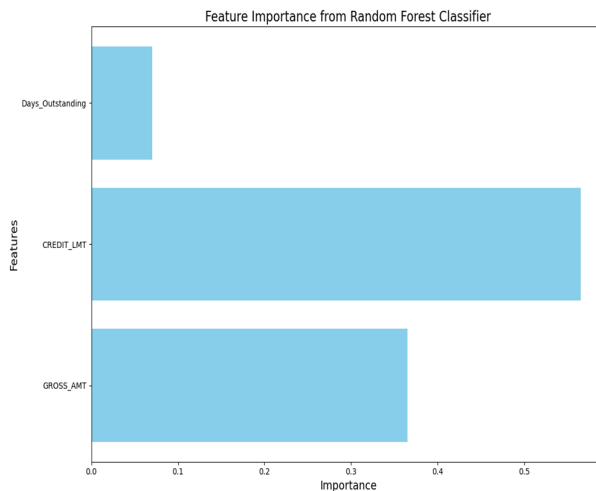

Fig. 7. Feature Importance Analysis

The results were instructive. The 'Credit Limit' feature, characterized by a score of 0.57, demonstrates the importance of the credit limit in signifying a company's financial bearings (David et al. 2002). The 'Gross Amount' followed suit with a score of 0.36, echoing Kou et al. (2019) regarding the indelibility of transactional volumes in decoding financial stability. Lastly, 'Days Outstanding', although registering a modest score of 0.07, reiterated its enduring significance, underpinning its centrality in financial assessments as postulated by Cho et al. (2020). These results are shown in Fig. 7.

3.4.8 Reporting & Feedback

In the final layer of the FTEDF, the emphasis is on the communication of detected anomalies to the auditing stakeholders. Reporting is not a mere presentation of errors, but a systematic exposition of irregularities accompanied by insights that drive corrective actions. This layer aids auditors in assimilating the nature and extent of errors, offering them a comprehensive understanding of the underlying issues. Such clarity can often expedite the audit process, making it both efficient and effective.

Table 4 presents a consolidated view of the insights that the FTEDF offers. Each row corresponds to a unique type of financial reporting discrepancy, its observed frequency, the potential root cause, and preliminary rectification steps. Such a table functions as an imperative tool, guiding auditors in their subsequent investigations and corrective measures. Further enriching the utility of this layer is the integration of feedback loops. These loops enable the continuous refinement of the FTEDF model, ensuring its adaptability to evolving transaction patterns and its resonance with the dynamic nature of business operations (Geerts 2011). The integration of feedback implies a two-way communication channel, allowing auditors to provide insights back into the model, ensuring that the model remains relevant and timely in its error detection capabilities. The iterative nature of this process underscores the commitment to accuracy and the pursuit of excellence in the realm of financial transaction audits.

Table 4
Insights for Users of FTEDF

Error Type	Frequency	Potential Root Cause	Preliminary Rectification Steps
Exceeding Credit Limit	High	Poor internal control mechanisms; aggressive revenue recognition strategies	Review and tighten credit control policies; enhance monitoring of accounts
Unmatched Transactions	Moderate	Discrepancies in client-master listing; manual entry errors	Cross-reference with source documents; provide training on accurate data entry
Invoice Volume Discrepancy	Low	Mismatch in record keeping; potential fraud or data tampering	Audit trail review; initiate a detailed transactional investigation
Delayed Days Outstanding	Moderate	Inefficient credit and collection procedures; potential liquidity issues	Review credit policy; enhance collection efforts and client follow-ups

In sum, the FTEDF introduces a transformative approach to assessing financial reporting quality. Rather than being confined to error detection, it encompasses multiple pivotal facets of financial data. This expansive purview ensures businesses are not just evaluated on their inaccuracies but also on the intricacies, patterns, and underlying contexts of their financial transactions. By weaving these dimensions into its analytic fabric, the FTEDF offers stakeholders a nuanced, layered, and robust understanding of a company's financial transparency, competency, and overall reporting integrity.

4. Discussion

The interplay between traditional accounting practice and machine learning stands at a pivotal juncture in the realm of AIS. As literature navigates this intricate milieu, there arises an imperative to not only innovate but also to articulate these findings to a diverse audience encompassing both academia and industry. Proper communication ensures the seamless transition of theoretical constructs into actionable insights, thus anchoring the research in practical relevance. We find that by incorporating the FRQS within the matrix of the Random Forest classifier, we can expand the error detection capabilities of AIS. Historically, as observed by David et al. (2002), AIS platforms have grappled with large and multifarious datasets, occasionally leading to potential discrepancies or oversights in identifying inconsistencies. Our reliance on machine learning, exemplified by the Random Forest classifier, signals a departure from these conventional challenges. Metrics, such as the 'Credit Limit' feature with a notable score of 0.57, bolster our claim of augmented accuracy, a direction also being pursued by Chan et al. (2018). Time, an irreplaceable commodity in audit processes, demands that AIS models optimize both speed and accuracy. Our research findings elucidate that the integration of machine learning within AIS can serve as a linchpin in bolstering the efficiency of error identification. The salience of this integration is underpinned by the classification matrix outcomes, highlighting machine learning's transformative potential in auditing. Sun (2019) has also championed the cause of integrating emerging technologies to reinvigorate conventional auditing techniques. One discernible limitation of our study was the non-inclusion of Objective 3, which sought to bridge the nexus between technological prowess and user-friendly interfaces. The pivotal role of a seamless interface, especially in a domain where end-users may not necessarily be technologically proficient, cannot be overstated. Geerts (2011) emphasizes this challenge, pointing out that the chasm between high-end technology and its tangible application in AIS often remains unbridged. By excluding this objective, our research may not address the broader

applicability of our findings across varied user bases. Future research endeavors could pivot around this dimension, striving to develop AIS solutions that are both potent and intuitively navigable.

Like all technological interventions, the application of ML in auditing is not without its hurdles. Brynjolfsson and Mitchell (2017) articulated some of these challenges, such as the conundrum of algorithm selection. With a myriad of algorithms available, each bringing its unique characteristics, biases, and limitations, auditors need to make judicious choices. A more pervasive challenge is the opacity of these algorithms. Dwivedi et al. (2021) discussed the 'black-box' nature of some sophisticated ML models, emphasizing that this could be a barrier to widespread adoption. In an industry where transparency is paramount and audit trails are sacrosanct, the enigmatic workings of these algorithms can be problematic. A salient feature of the contemporary financial landscape is its inherent dynamism, marked by shifting regulations, novel business methodologies, and swift technological evolutions. A successful AIS paradigm should, thus, be imbued with the qualities of adaptability and scalability. This ensures its relevance even as businesses grapple with burgeoning transaction volumes and intricacies. This ethos of constant evolution and adaptability is rooted by Alles et al. (2021), who accentuate the urgency of crafting AIS models that can withstand the test of time. The essence of machine learning, particularly in the context of AIS, lies in its capacity for continuous learning and refinement. As our AIS model is inundated with newer datasets, its algorithms should adapt, evolve, and enhance their precision. By facilitating this continuous learning mechanism, we hope to future-proof our AIS model, ensuring it remains at the vanguard of error detection, regardless of emergent transactional patterns. This paradigm of ceaseless improvement finds echoes in Holzinger et al. (2018), who have spotlighted the imperativeness of adaptability in machine learning frameworks. These findings offer a vision of the AIS landscape's potential trajectory, a melding of time-tested accounting practices with the dynamism of machine learning. Through this discourse, we aim not just to proffer a novel AIS model but to stimulate robust conversations within the AIS community, promoting a culture of collective insights and mutual evolution.

5. Conclusion

In the constantly evolving nexus of accounting and technological innovation, AIS stand out as pivotal instruments in redefining traditional accounting procedures. This research is grounded in the merger of machine learning paradigms with AIS, offering insights that traverse both academic profundity and practical applicability. Central to the DSM in AIS research is the meticulous development of a framework tailored to address distinct issues flagged in the earlier phases of the study. This structured approach facilitates AIS researchers to resonate their strategies with grander visions, thereby sculpting tools or systems that adeptly tackle real-world predicaments. As postulated by Gregor and Hevner (2013), a judiciously crafted framework serves a dual purpose: it offers an architectural design for the research and doubles as a schematic for the envisaged solution's execution. Within this study's purview, the framework orbits around the timely and essential task of detecting financial transaction errors—a burgeoning concern in the current AIS context. As the global financial realm undergoes rapid metamorphosis and as we grapple with the tidal wave of transaction data spawned daily, the clarion call for a nimble, progressive, and responsive solution becomes increasingly loud. Considering this exigency, we unfurl the FTEDF. Envisioned as a stratified, machine learning-centric system, the FTEDF is tailored for both precision and efficiency in identifying anomalies in financial transactions. Amidst the seismic shifts in regulatory dynamics, novel business tactics, and swift technological advancements, an AIS's utility hinges on its ability to adapt and scale. Our investigative foray, synergizing with the insights of Alles et al. (2021), emphasizes the pressing need for AIS models that are both visionary and versatile, ready to confront the morrow's challenges.

While this research has made significant strides in advancing our understanding of error detection within AIS through the FTEDF, it is imperative to acknowledge its limitations for a comprehensive perspective. The omission of Objective 3, focusing on user-interface nuances, is a primary limitation in our study. Our reliance on symmetric data may not capture the complexities of real-world financial transactions, and the exclusive use of the Random Forest classifier means we might not harness the full potential of machine learning tools available for AIS research. These constraints suggest areas for future refinements in the framework. Drawing upon Geerts (2011), the confluence of technological prowess and user-centered design emerges as crucial, especially in an arena frequented by a diverse array of users. It is an avenue ripe for exploration by future AIS. In retrospection, AIS emerges not as a monolithic entity but a dynamic mosaic of continual innovation and growth. Our machine learning-infused model exemplifies this spirit of perpetual refinement. Echoing Cho et al. (2020), AIS's future trajectory must tread the path of proactive evolution, staying one step ahead of the curve. This endeavor is more than an innovative AIS paradigm. It extends an invitation to the larger AIS fraternity to envision a future where time-honored practices meld seamlessly with trailblazing innovations, carving out avenues that are transformative and inclusive. Our aspiration is for this research to ignite deeper conversations, catalyzing further in-depth explorations into the boundless domain of AIS.

Author Contributions

Conceptualization, E.A. and A.K.; methodology, E.A. and A.K.; formal analysis, E.A.; writing—original draft preparation, E.A.; writing—review and editing, K.G. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement

The data presented in this study can be made available upon request from the authors.

References

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.

Alles, M. G., Dai, J., & Vasarhelyi, M. A. (2021). Reporting 4.0: Business reporting for the age of mass customization. *Journal of Emerging Technologies in Accounting*, *18*(1), 1-15.

An, B., & Suh, Y. (2020). Identifying financial statement fraud with decision rules obtained from Modified Random Forest. *Data Technologies and Applications*, *54*(2), 235-255.

Barr-Pulliam, D., Brown-Liburd, H. L., & Sanderson, K. A. (2022). The effects of the internal control opinion and use of audit data analytics on perceptions of audit quality, assurance, and auditor negligence. *Auditing: A Journal of Practice & Theory*, *41*(1), 25-48.

Beven, K., & Binley, A. (1992). The future of distributed models: model calibration and uncertainty prediction. *Hydrological processes*, *6*(3), 279-298.

Breiman, L. (2001). Random forests. *Machine learning*, *45*, 5-32.

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.

Brynjolfsson, E., & Mitchell, T. (2017). What can machine learning do? Workforce implications. *Science*, *358*(6370), 1530-1534.

Byrnes, P. E., Al-Awadhi, A., Gullvist, B., Brown-Liburd, H., Teeter, R., Warren Jr, J. D., & Vasarhelyi, M. (2018). Evolution of auditing: From the traditional approach to the future audit. In *Continuous auditing: Theory and application* (pp. 285-297). Emerald Publishing Limited.

Canhoto, A. I., & Clear, F. (2020). Artificial intelligence and machine learning as business tools: A framework for diagnosing value destruction potential. *Business Horizons*, *63*(2), 183-193.

Cho, S., Vasarhelyi, M. A., Sun, T., & Zhang, C. (2020). Learning from machine learning in accounting and assurance. *Journal of Emerging Technologies in Accounting*, *17*(1), 1-10.

Cohen, J., Ding, Y., Lesage, C., & Stolowy, H. (2012). *Corporate fraud and managers' behavior: Evidence from the press* (pp. 271-315). Springer Netherlands.

Costa, M., Lisboa, I., & Gameiro, A. (2022). Is the Financial Report Quality Important in the Default Prediction? SME Portuguese Construction Sector Evidence. *Risks*, *10*(5), 98.

David, J. S., Gerard, G. J., & McCarthy, W. E. (2002). Design science: building the future of AIS. *American Accounting Association*, *69*.

Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., ... & Williams, M. D. (2021). Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management*, *57*, 101994.

Geerts, G. L. (2011). A design science research methodology and its application to accounting information systems research. *International journal of accounting Information Systems*, *12*(2), 142-151.

Gregor, S., & Hevner, A. R. (2013). Positioning and presenting design science research for maximum impact. *MIS quarterly*, *37*(2), 337-355.

Groomer, S. M., & Murthy, U. S. (2018). Continuous Auditing of Database Applications: An Embedded Audit Module Approach1. In *Continuous auditing* (pp. 105-124). Emerald Publishing Limited.

Healy, P. M., & Palepu, K. G. (2003). The fall of Enron. *Journal of economic perspectives*, *17*(2), 3-26.

Hevner, A. R., March, S. T., Park, J., Ram, S., & Ram, S. (2004). Research essay design science in information. *MIS Q*, *28*(1), 75-105.

Holzinger, A., Kieseberg, P., Weippl, E., & Tjoa, A. M. (2018). Current advances, trends and challenges of machine learning and knowledge extraction: from machine learning to explainable AI. In *Machine Learning and Knowledge Extraction: Second IFIP TC 5, TC 8/WG 8.4, 8.9, TC 12/WG 12.9 International Cross-Domain Conference, CD-MAKE 2018, Hamburg, Germany, August 27–30, 2018, Proceedings 2* (pp. 1-8). Springer International Publishing.

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.

Kou, G., Chao, X., Peng, Y., Alsaadi, F. E., & Herrera Viedma, E. (2019). Machine learning methods for systemic risk analysis in financial sectors.

Romney, M., Steinbart, P., Mula, J., McNamara, R., & Tonkin, T. (2012). *Accounting Information Systems Australasian Edition*. Pearson Higher Education AU.

Roussy, M., Barbe, O., & Raimbault, S. (2020). Internal audit: from effectiveness to organizational significance. *Managerial Auditing Journal*, *35*(2), 322-342.

Stoel, D., Havelka, D., & Merhout, J. W. (2012). An analysis of attributes that impact information technology audit quality: A study of IT and financial audit practitioners. *International Journal of Accounting Information Systems*, *13*(1), 60-79.

Studer, S., Bui, T. B., Drescher, C., Hanuschkin, A., Winkler, L., Peters, S., & Müller, K. R. (2021). Towards CRISP-ML (Q): a machine learning process model with quality assurance methodology. *Machine learning and knowledge extraction*, *3*(2), 392-413.

Sun, T. (2019). Applying deep learning to audit procedures: An illustrative framework. *Accounting Horizons*, *33*(3), 89-109.

Turkay, C., Laramee, R., & Holzinger, A. (2017). On the challenges and opportunities in visualization for machine learning and knowledge extraction: A research agenda. In Machine Learning and Knowledge Extraction: First IFIP TC 5, WG 8.4, 8.9, 12.9 International Cross-Domain Conference, CD-MAKE 2017, Reggio, Italy, August 29–September 1, 2017, Proceedings 1 (pp. 191-198). Springer International Publishing.

Vasarhelyi, M. A., Halper, F. B., & Ezawa, K. J. (1991). The continuous process audit system: A UNIX-based auditing tool. *The EDP Auditor Journal*, *3*(3), 85-91.

Vasarhelyi, M. A., Kogan, A., & Tuttle, B. M. (2015). Big data in accounting: An overview. *Accounting Horizons*, *29*(2), 381-396.

1764