

Explainable Anomaly Detection in Hospitality Revenue Records: A Scenario-Based Modeling Study

Dominick Aondakaa^{1,*}, Joy Kweku Sakyi², Ann Kinyanjui³, Esther Ogechi Nwozaku⁴

¹Senior Transaction Monitoring Analyst, Coinbase, United States

²Independent Researcher, Rock Hill, South Carolina, United States

³Finance Manager, Amazon, Seattle,
Washington, United States

⁴Cage Supervisor, MGM National Harbor
Hotel & Casino, Maryland, United States

*Corresponding author: Dominick Aondakaa
(aondakaadominick@gmail.com)

Email addresses: aondakaadominick@gmail.com
(D. Aondakaa); joy.sakyi@outlook.com (J. K. Sakyi);

awk75@cornell.edu (A. Kinyanjui);
nwozakuesther@gmail.com (E. O. Nwozaku)

[https://doi.org/DOI: 10.56201/jhmtr.vol.6no2.2025.pg58.94](https://doi.org/DOI:10.56201/jhmtr.vol.6no2.2025.pg58.94)

Abstract

Hotel revenue records are produced across reservation, property management and point-of-sale systems, and small irregularities in refunds, allowances and repeated postings can quietly erode revenue or conceal misconduct. Detection tools help only if reviewers can see why a record was flagged and if alert volumes fit the time a finance team actually has. A scenario-based model of one year of activity in four hotels compared simple policy rules, an expert points scorecard and interpretable statistical models against an opaque benchmark, with each model developed on early 2025 data and tested on later, unseen months. Conventional exception rules caught most irregularities but buried them among false alerts, whereas a sparse logistic model built on indicators a reviewer can verify on the guest folio found most anomalies, with more than four in five alerts genuine and less than an hour of weekly review per hotel. Unusual refunds and adjustments were readily detectable, but duplicate postings remained a persistent weakness when point-of-sale identifiers were missing. Performance depended as much on governance as on model choice: fixed thresholds lost precision as irregularities became rarer, peer baselines broke after price changes, and duplicate checks failed under group billing. A two-tier review queue that placed high-confidence scorecard alerts ahead of model alerts found about three quarters of all anomalies at roughly one review hour per hotel per week.

Keywords: anomaly detection; explainable models; hospitality accounting; revenue assurance; duplicate transactions; class imbalance; reviewer workload

1. Introduction

Revenue records in a hotel are produced by many hands. Night auditors post room and tax charges, outlet cashiers post food and beverage checks, front-desk agents issue refunds and service-

recovery allowances, and accounting staff make corrections after guests leave. Each of these activities is legitimate in most cases, and each also offers a route for error or abuse. Occupational fraud studies consistently identify billing schemes, false refunds, and voided or adjusted sales among the common schemes in cash-handling businesses, and they report that a large share of cases are found by tip or by accident rather than by monitoring (ACFE, 2024). In service-intensive businesses such as hotels, these losses fall under revenue assurance, where analytics are increasingly used to find leakage between what was sold and what was collected (Pierce et al., 2015; Sakyi et al., 2023; Sanni & Wedraogo, 2024). Duplicate postings, which are usually honest mistakes, cause guest disputes, chargebacks, and reputational damage when they reach a guest folio. Most hotels rely on exception reports built into their property management systems: lists of all refunds above a limit, all allowances without manager approval, and all voids by user. These reports are transparent, but they are long, and reviewers learn to skim them. Machine learning methods promise shorter and better-ranked lists (Atakpa & Abetoh, 2022; Bello et al., 2025a; Dal Pozzolo et al., 2014), yet many are difficult to explain to an auditor, a general manager, or an employee whose transaction was questioned. In accounting control settings, an alert that cannot be explained is hard to act on and hard to defend, and auditors of AI-supported financial systems increasingly require evidence that a human can validate (Barredo Arrieta et al., 2020; Elebe & Anunagba, 2025; Fobellah, 2025b; Rudin, 2019) (Nesvijejskaia et al., 2021).

This paper asks which transparent methods identify unusual refunds, adjustments, and duplicate entries with review burdens that a hotel finance team can carry. It answers the question with a controlled simulation, because real hotel ledgers with verified anomaly labels are rarely available for research and because a simulation lets us vary the conditions under which a method might fail. The study follows the ADEMP structure for simulation studies (aims, data-generating mechanisms, estimands, methods, performance measures) recommended by Morris et al. (2019), which builds on earlier guidance for the design of simulation studies (Burton et al., 2006).

Prior work leaves three gaps. Hospitality research has examined revenue generation in depth but has given little attention to transaction-level revenue integrity, and the few hotel anomaly studies do not report review burden (Ivanov & Zhechev, 2012; Mendes et al., 2022; Wibawa et al., 2025). Fraud analytics studies, including many recent conceptual frameworks, seldom evaluate transparent methods on held-out data with explicit workload measures (Fadayomi et al., 2024; Hilal et al., 2022; Ominyi et al., 2024). Guidance on class imbalance and calibration, although well developed, is rarely translated into operating rules for finance teams (Dal Pozzolo et al., 2015; Van Calster et al., 2019). This study addresses these gaps. It introduces a reproducible scenario generator for hotel revenue ledgers that contains ten anomaly types and, equally important, legitimate look-alike patterns such as repeat bar orders, service-recovery comps, goodwill credits, late dispute refunds, and group master folios. It compares rule-based and interpretable statistical detectors on development and evaluation data separated in time, with every threshold fixed before evaluation. It reports reviewer workload beside accuracy at both a best-accuracy and a capacity-constrained threshold, and it stress-tests each method against class imbalance and changing transaction patterns. We are not aware of a prior study that evaluates transparent detectors of hotel refund, adjustment, and duplicate anomalies jointly on accuracy and reviewer workload, and the two-tier review model proposed here is evaluated on held-out data rather than described conceptually. The rest of this section reviews related work on revenue integrity in hospitality, on fraud, audit, and anomaly detection, on imbalanced learning and interpretability, and on data governance and operational analytics. Section 2 describes the research method, Section 3 presents and discusses the results, and Section 4 concludes.

1.1 Revenue integrity in hospitality

Hotels earn revenue through a chain of operational systems: reservations and rate management upstream, and the property management system (PMS), point-of-sale (POS) terminals, and night audit downstream. The revenue management literature has concentrated on the upstream chain, showing how pricing, forecasting, and capacity control raise revenue in perishable-inventory businesses (Cross et al., 2009; Ivanov & Zhechev, 2012; Kimes, 1989; Talluri & van Ryzin, 2004; Weatherford & Kimes, 2003). Booking-level data are now openly available for this purpose, and cancellation prediction has become a standard analytic task (Antonio et al., 2017, 2019). Information technology adoption in hotels has been linked to operating performance (Ham et al., 2005; Law et al., 2009), and restaurant revenue management has been extended to POS-level data (Kimes, 2008). The downstream question this paper addresses, whether the revenue that was earned is the revenue that was recorded and kept, has received far less attention.

Hotel accounting conventions define where refunds, allowances, and corrections belong in the revenue accounts. The Uniform System of Accounts for the Lodging Industry, revised in its 12th edition, standardizes how allowances are netted against departmental revenue (HFTP, 2024). Empirical work on irregularities in hospitality is sparse but consistent. Surveys of food service operators found employee theft and policy violations to be common and controls to vary in effectiveness (Ghiselli & Ismail, 1995). A study of star-rated hotels in Bandung linked internal audit and unethical behavior to accounting fraud, while the formal effectiveness of internal control did not show the expected effect (Julianty et al., 2022). A case study of cash receipt fraud at a Yogyakarta hotel traced loss to weak monitoring despite documented procedures (Enshara & Bastian, 2018). The strongest causal evidence comes from restaurants, where installing theft-monitoring software across 392 locations reduced theft and raised productivity (Pierce et al., 2015). Anomaly detection has begun to appear in hotel settings, for utility consumption in hotel units (Mendes et al., 2022) and for hotel transaction records in an information-system prototype (Wibawa et al., 2025), but these studies do not report reviewer workload or test robustness to business changes. Outside hospitality, revenue assurance has been framed as a data problem in service-intensive organizations, where analytics reconcile what was sold with what was billed and collected (Sakya et al., 2023; Sanni & Wedraogo, 2024). Related work describes revenue optimization through integrated planning (Nnabueze et al., 2024), customer service analytics as a revenue driver (Sakya et al., 2022a), and automation of utility and service operations with analytics embedded in the workflow (Nnabueze et al., 2023). Hospital revenue-cycle studies are a close analogue: they address billing, claims, and cost transparency in organizations with high transaction volumes and many hands posting charges (Odeyemi & Idowu, 2025). Together these strands suggest that hotels need transaction-level controls that complement upstream revenue management, which motivates the present study.

1.2 Fraud, audit, and anomaly detection

The fraud triangle of pressure, opportunity, and rationalization remains the dominant explanation for occupational fraud (Cressey, 1953), and the hotel case evidence cited above fits it closely (Enshara & Bastian, 2018). Global occupational fraud surveys report that billing and refund schemes are common in cash-handling sectors and that many frauds are found by tips rather than by monitoring (ACFE, 2024). Internal control and fraud risk frameworks place data analytics at the center of detective controls (COSO & ACFE, 2023; COSO, 2013). Conceptual models from the internal audit literature propose technology-enabled and risk-based approaches to audit quality and internal control design (Akomolafe & Agu, 2018, 2019a; Akomolafe et al., 2023a), data-driven

risk evaluation and compliance-risk measurement in emerging markets (Akomolafe & Agu, 2019b; Akomolafe et al., 2023b), and governance strategies for financial resilience (Akomolafe & Agu, 2019c). More recent work extends these models to data-driven compliance analytics and Python-based predictive risk assessment (Agu et al., 2025; Akomolafe et al., 2025), and reviews SOX compliance frameworks for cross-border audits (Agu et al., 2023).

Audit research has argued for two decades that analytics should move from periodic sampling to continuous, data-level testing (Alles et al., 2008; Kogan et al., 2014). Big data has been described as changing both the scope of accounting evidence and the skills auditors need (Appelbaum et al., 2017; Vasarhelyi et al., 2015), and artificial intelligence is expected to formalize parts of audit judgment and supplement the audit workforce (Issa et al., 2016; Kokina & Davenport, 2017). Recent conceptual models describe AI-enabled IT general controls and SOX automation (Mbonu et al., 2022a), AI-assisted continuous auditing for technology risk (Aliliele et al., 2025a), analytics to strengthen audit quality in financial institutions (Ozowara et al., 2025), and audit analytics for accountability in grant-funded programs (Abetoh & Atakpa, 2024). Other work addresses audit liaison and corrective-action planning (Ebhojie et al., 2023), tamper-resistant recordkeeping for audit readiness (Anichukwueze et al., 2021; Oshoba et al., 2020), real-time compliance pipelines (Uchechi et al., 2022), and the standardization of compliance controls in high-volume, decentralized networks (Uchechi et al., 2023), which resembles the multi-outlet structure of a hotel. Strategic accounting systems and financial governance in smaller enterprises have been linked to transparency and compliance (Bola-Sadipe et al., 2019; Eboh et al., 2025). At the same time, the cost of compliance is real: regulation can create burdens that organizations struggle to absorb (Fobellah, 2025c), and auditors themselves are subject to cognitive biases that structured tools can help counter (Fobellah, 2025a). These points support the emphasis on workload and on clear reasons for each alert in the present design. Surveys of statistical fraud detection distinguish supervised methods, which learn from confirmed cases, from unsupervised methods that score how unusual a record is (Abdallah et al., 2016; Bolton & Hand, 2002; Ngai et al., 2011; Phua et al., 2010; West & Bhattacharya, 2016). Practitioner-focused work on card fraud has documented the realities that shape any deployed detector: extreme class imbalance, delayed and incomplete labels, concept drift, and the limited number of alerts investigators can review (Dal Pozzolo et al., 2014, 2018). Comparative studies show that relatively simple models remain competitive (Bhattacharyya et al., 2011), and hybrid designs that combine unsupervised scores with supervised learners can improve detection (Carcillo et al., 2021). Book-length treatments synthesize these lessons for fraud analytics teams (Baesens et al., 2015), and a recent review maps anomaly detection techniques across financial fraud types (Hilal et al., 2022).

Accounting data have their own detection traditions. Digit-frequency tests based on Benford's law are an established analytical procedure (Nigrini & Mittermaier, 1997; Nigrini, 2012). Studies of financial statement fraud compare statistical and machine learning classifiers (Bao et al., 2020; Perols, 2011), while journal-entry mining, cluster analysis, and deep autoencoders have been applied to ledger records directly (Debreceeny & Gray, 2010; Schreyer et al., 2017; Thiprungsri & Vasarhelyi, 2011). General anomaly detection surveys cover the methods used across domains (Chandola et al., 2009; Pang et al., 2021), and robust statistics based on the median and median absolute deviation provide transparent outlier screens (Iglewicz & Hoaglin, 1993; Leys et al., 2013; Rousseeuw & Hubert, 2011). Isolation forests are a widely used unsupervised baseline (Liu et al., 2008, 2012), and benchmark studies show that the ranking of unsupervised detectors depends strongly on the dataset and evaluation measure (Campos et al., 2016; Goldstein & Uchida, 2016).

A growing body of conceptual and review work applies these methods to particular sectors. Reviews cover machine learning advances for banking fraud (Atakpa & Abetoh, 2022), predictive analytics for anti-money laundering (Atakpa, 2021), AI for healthcare payment fraud (Atakpa et al., 2024), and fraud detection in UK digital payment systems (Bello et al., 2025a). Conceptual models propose anomaly detection for financial time series (Atakpa & Fobellah, 2023), AI-assisted compliance monitoring and anomaly detection in high-volume transaction environments (Ominyi et al., 2024), adaptive risk scoring for real-time transaction monitoring (Fadayomi et al., 2024), and predictive analytics for fraud prevention in public systems (Adelanwa et al., 2023). Related proposals address regulatory technology for fraud reporting accuracy (Bello et al., 2025b), analyst decision support in financial crime investigation (Okoruwa et al., 2025), AI risk scoring and fraud detection for cross-border payments (Adesuyi et al., 2024, 2025), end-to-end monitoring controls for correspondent banking (Okafor et al., 2024), and blockchain support for know-your-customer and anti-money laundering checks (Bello et al., 2025c). Forensic analytics have been described for cyber-enabled fraud in telemedicine and for voice-over-IP communications (Ekechi et al., 2022; Mbonu et al., 2021b), and risk analytics for investment fraud and FinTech monitoring (Eboh & Aliliele, 2023, 2024, 2025). Most of these contributions are architectural; few report precision, recall, and workload on held-out data, which is the gap this study addresses for hotel ledgers.

1.3 Imbalanced learning, evaluation, and interpretability

Rare-event learning has a long literature. Logistic regression is known to underestimate rare-event probabilities without correction (King & Zeng, 2001); resampling methods such as SMOTE and undersampling, and cost-sensitive learning, are the common remedies (Chawla et al., 2002; Elkan, 2001; He & Garcia, 2009). Reviews of imbalanced learning catalogue hundreds of variants and stress that their benefit depends on the evaluation criterion (Haixiang et al., 2017; Krawczyk, 2016). Precision-recall analysis is more informative than ROC analysis when positives are rare (Davis & Goadrich, 2006; Saito & Rehmsmeier, 2015), although ROC concepts remain useful for reasoning about operating points under changing costs and class distributions (Fawcett, 2006; Provost & Fawcett, 2001). Resampling and weighting also change predicted probabilities, so calibration and threshold choice must be revisited after any such step (Dal Pozzolo et al., 2015; Niculescu-Mizil & Caruana, 2005; Van Calster et al., 2019). When data change over time, models face dataset shift and concept drift, which must be detected and managed rather than assumed away (Gama et al., 2014; Moreno-Torres et al., 2012). Leakage from future information is a common cause of optimistic evaluation, which argues for strict temporal separation of development and evaluation data (Kaufman et al., 2012). Resampling-based uncertainty for dependent data requires block or stationary bootstrap schemes rather than independent resampling of records (Efron & Tibshirani, 1993; Künsch, 1989; Politis & Romano, 1994). Interpretable models are those whose reasoning can be inspected directly, as opposed to black-box models explained after the fact (Doshi-Velez & Kim, 2017; Lipton, 2018; Rudin, 2019). Post-hoc explanation tools such as local surrogate models and Shapley-value attributions are widely used (Lundberg & Lee, 2017; Ribeiro et al., 2016), and surveys catalogue their strengths and limits (Barredo Arrieta et al., 2020; Guidotti et al., 2018). Social-science work on explanation shows that people prefer contrastive, selective, and causal reasons (Miller, 2019), and that they abandon algorithms after seeing them err unless they understand them (Dietvorst et al., 2015). Sparse linear models, scorecards, and generalized additive models can match complex models on structured data while remaining readable (Breiman, 2001; Caruana et al., 2015; Tibshirani, 1996; Ustun & Rudin, 2016); scorecards in particular have

decades of use in credit decisions (Siddiqi, 2006). Explainability has been examined specifically for credit risk and for fraud detection (Bussmann et al., 2021; Psychoula et al., 2021).

Governance literature adds that automated decisions in regulated settings need accountability, auditability, and human oversight (Olagunju et al., 2023). Work in this area covers algorithmic accountability and model risk management in financial services (Ominyi & Anichukwueze, 2023), responsible governance structures for automated decision systems (Ominyi & Anichukwueze, 2024), accountability and trade-secret tensions in AI (Annan, 2024), anti-discrimination compliance for automated decisions (Annan, 2025), and release-gating practices for model governance (Nwakamma et al., 2024). Human-in-the-loop machine learning, in which reviewers' decisions feed back into models, is directly relevant to a review queue whose outcomes become future labels (Ladapo et al., 2022, 2024). Responsible AI in IT auditing emphasizes evidence validation and human oversight (Elebe & Anunagba, 2025), auditing AI-powered financial systems raises similar concerns (Fobellah, 2025b), and interpretable machine learning has been applied in other operational domains for the same reasons (Komi & Adamolekun, 2021). Process-mining and adaptive-control approaches show how automated workflows can be made auditable in compliance-sensitive environments (Sanni et al., 2022, 2025).

Simulation studies allow methods to be compared when the truth is known and conditions can be varied systematically. Guidance on their design emphasizes stating aims, data-generating mechanisms, estimands, methods, and performance measures in advance (Burton et al., 2006; Morris et al., 2019), and simulation model credibility rests on verification and validation (Sargent, 2013). In fraud research, agent-based payment simulators were developed precisely because labeled transaction data cannot be shared (Lopez-Rojas & Axelsson, 2014; Lopez-Rojas et al., 2016). Synthetic data are therefore a recognized research instrument, but conclusions drawn from them should focus on relative performance and failure modes rather than absolute accuracy.

1.4 Data governance and operational analytics

Detection is only as good as the data extract it runs on. Work on data governance addresses sensitivity classification and regulatory traceability (Aliliele et al., 2024a), lakehouse governance for loss prevention and compliance (Aliliele et al., 2025b), privacy-preserving architectures for regulated analytics (Atakpa & Abolaji, 2022), data protection impact assessment in multi-cloud financial systems (Mbonu et al., 2022b), and the harmonization of privacy regimes across jurisdictions (Ominyi & Anichukwueze, 2022). Risk-based business intelligence architectures and API governance frameworks for financial platforms describe how transaction data move between systems (Aliliele et al., 2023; Mbonu et al., 2021a), and ETL design for integrated enterprise data describes how they are joined (Badmus et al., 2019). Log analytics, security analytics, and digital forensics provide the evidentiary trail behind an alert (Mbonu et al., 2019; Ogbole et al., 2025). National-scale data programs show that interoperability and data quality must be designed in, not assumed (Ilodigwe & Adesemoye, 2021), and standardized reporting workflows improve both data quality and management visibility (Alieme et al., 2022). The duplicate-key failure reported later in this paper is, at bottom, a data governance finding. Finance and operations research has produced many models for turning transaction data into decisions. These include data-led cost governance and predictive budgeting (Adesuyi et al., 2022, 2023), variance analytics for cost control (Aliliele et al., 2024b), automated financial consolidation (Bola-Sadipe et al., 2023), financial reporting for compliance and accountability (Medon & Oduleye, 2022), predictive forecasting and cash-flow models (Medon & Oduleye, 2024; Oduleye & Medon, 2023; Tonoyan et al., 2021b), data-driven cost management (Oduleye & Medon, 2021), executive decision systems and dashboards (Lawal

& Oduleye, 2019; Walawalkar et al., 2025), demand forecasting for revenue projection (Tonoyan et al., 2021a), and scenario-based financial modelling for planning (Filani et al., 2023). Simulation modelling has been used to estimate the financial impact of process changes (Tonoyan et al., 2022), and real-time KPI tracking has been reviewed as a monitoring mechanism (Tonoyan et al., 2024). Operations research on service businesses adds the organizational side. Digital transformation of operations combines automation, analytics, and process redesign (Agboada et al., 2018), standard operating procedures act as infrastructure for sustaining improvement (Agboada et al., 2019), and Lean Six Sigma has been adapted for smaller organizations (Ambali et al., 2021). Performance dashboards, KPI frameworks, and accountability metrics give leaders visibility of operations (Edivri & Oteri, 2024; Ezeako et al., 2024; Sakyi et al., 2022b; Seyi-Lande et al., 2022). Multi-site operational efficiency, governance across retail networks, and the use of real-time operational and workforce data for customer experience are especially relevant to multi-property hotel groups (Akhigbe et al., 2023; Akin-Oluyomi & Akhigbe, 2022, 2023). Procurement studies show how process automation, digital platforms, and visualization improve transparency in transaction-heavy functions (Akinleye & Adeyoyin, 2021; Babatope et al., 2023; Okoruwa et al., 2024), and churn analytics in telecommunications show how predictive models are embedded in service operations (Oziri et al., 2020). The present study borrows from this tradition the view that an analytic is useful only if it fits the time and attention of the people who must act on it.

2. Research Method

2.1 Study design

Table 1 summarizes the design in ADEMP form. Figure 1 shows how data flowed from the generator to the outcome measures.

Table 1. Study design summary (ADEMP).

Element	Specification
Aims	Compare transparent detectors of unusual refunds, adjustments, and duplicate entries on accuracy and reviewer workload; examine class imbalance and sensitivity to transaction patterns.
Data-generating mechanism	Four hotels (90 to 200 rooms), daily arrivals with seasonality, Jan 1 to Dec 31, 2025; legitimate charges, credits, and look-alike patterns; ten injected anomaly types at 0.5% of charges (baseline).
Estimands	Precision, recall, false-positive rate, F1, area under the precision-recall curve (PR-AUC), alerts and review hours per property-week, alerts per true positive.
Methods	R1 policy rule set; R2 expert scorecard; S1 peer robust z-score; S2 L1 logistic regression; S3 depth-4 decision tree; isolation forest as opaque reference.
Performance measurement	Fit Jan to Jun; calibrate Jul to Aug; evaluate Sep to Dec. Block bootstrap (300 resamples of property-weeks); three replicate datasets; seven evaluation scenarios.

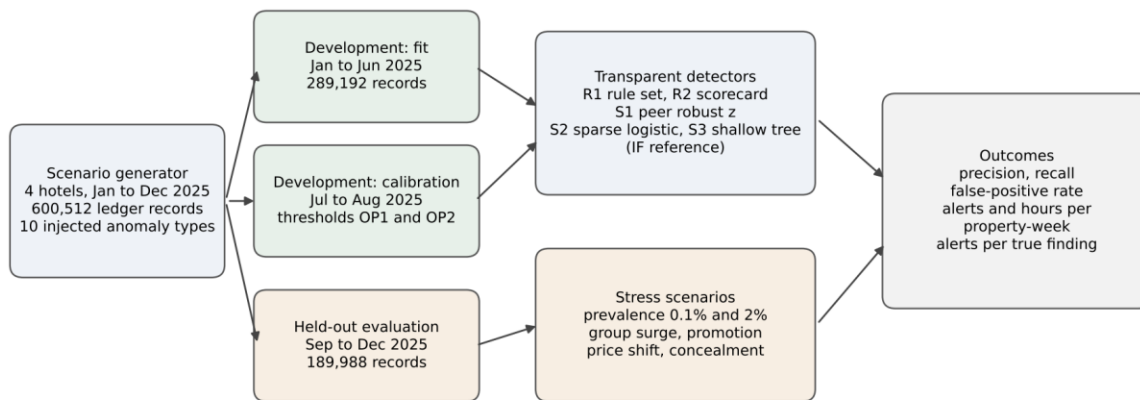


Figure 1. Study design. Development data (fit and calibration) and evaluation data are separated in time; stress scenarios replace the evaluation period only.

2.2 Data-generating mechanism

Properties and stays. Four properties with 160, 120, 90, and 200 rooms and average daily rates of 145, 110, 85, and 175 currency units were simulated for every day of 2025. Daily arrivals followed a Poisson distribution with a seasonal mean (peak in spring and summer, about 72% average occupancy) and a length of stay of one to ten nights. Each folio received a nightly room charge posted by the night audit between 02:30 and 03:10, an optional daily breakfast package at a fixed price, restaurant, bar, and room-service checks, and occasional spa, minibar, laundry, and parking charges. Bar checks were drawn from a fixed menu so that identical amounts occur naturally, and 8% of bar checks were followed by a second identical round within 2 to 25 minutes. Each property had 14 front-office and outlet clerks plus an audit user.

Legitimate credits and look-alikes. About 4% of folios received an allowance against a charge; 12% of those were service-recovery comps of 50% to 100% of the charge. Allowances of 50 or more required approval, which was recorded 92% of the time, and 6% of allowances were posted between midnight and 06:00 as audit corrections. About 3% of folios received a refund, of which 6% were dispute refunds 8 to 30 days after checkout and 6% were unlinked deposit refunds. A further 0.8% of folios received an unlinked loyalty goodwill credit in a round amount. These patterns were included on purpose: each resembles one of the anomaly types and is a realistic source of false alerts.

Anomalies. Anomalies were injected at a rate of 0.5% of charge records in the development period and at the scenario rate in the evaluation period. Each anomaly was attached to a randomly chosen host charge and assigned to one of three families and ten types (Table 2). For refunds and adjustments, 60% were posted by one of two clerks per property and 35% were posted between midnight and 05:00, reflecting the concentration of misconduct in a few users and outside supervised hours. Duplicates kept the clerk of the original record.

Table 2. Anomaly taxonomy and counts.

Family and type	Definition	All data	Evaluation
Refund: over-refund	Refund 1.2 to 2.5 times the linked charge	312	99
Refund: unlinked	Refund with no matching charge	307	93
Refund: late	Refund 10 to 60 days after checkout	176	59
Refund: repeat	Second partial refund taking the total above the charge	253	77
Adjustment: large share	Allowance of 50% to 100% of a charge, usually unapproved	427	116
Adjustment: threshold hugging	Allowance of 42.00 to 49.99, just below the approval limit	369	125
Adjustment: round unlinked	Credit of 50, 75, 100, 150, or 200 with no linked charge	262	78
Duplicate: exact	Same charge reposted within 6 minutes	385	121
Duplicate: delayed	Same charge reposted 1 to 30 hours later	405	147
Duplicate: room night	Second room charge for the same night	61	22
Total		2,957	937

Evaluation counts refer to September to December 2025. Anomalies make up 0.49% of all records.

2.3 Separation of development and evaluation data

Records were split by posting date. Records from January to June (289,192 records, 1,408 anomalies) were used to fit models and to compute peer-group medians and median absolute deviations. Records from July to August (118,649 records, 576 anomalies) were used only to choose thresholds. Records from September to December (189,988 records, 937 anomalies, 72 property-weeks) were used only for evaluation. Credits posted after December 31, 2025 for December stays (2,683 records) were excluded. All features were computed from information available at posting time: duplicate flags look only at earlier records, and clerk activity uses the previous 30 days. Nothing estimated on evaluation data was used by any detector.

2.4 Transparent indicators

All detectors draw on the same 17 indicators, each of which a reviewer can check on the folio itself (Table 3). One candidate indicator, a flag for any repeat of the same item and amount on a folio within 36 hours, was dropped during development because it fired on 32.3% of normal records: nightly room charges and daily breakfast packages repeat by design. It was replaced by two narrower flags (repeat on the same business date; repeat of a non-recurring item). This is a first, small illustration of sensitivity to transaction patterns, and of the data-quality point that a key must reflect how records are actually produced (Alieme et al., 2022; Ilodigwe & Adesemoye, 2021).

Table 3. Indicators available to all detectors.

Indicator	Definition
Record is a refund / adjustment	Transaction type flags
Credit ratio	Credit amount divided by the linked charge
Cumulative credit ratio	All credits so far against the linked charge divided by that charge
Unlinked credit	Credit with no linked charge
Days after checkout	Posting time minus checkout time, for credits
Off-hours credit	Credit posted between 00:00 and 06:00
Unapproved credit at or above limit	Credit of 50 or more with no approval code
Credit just under limit	Credit between 42.50 and 49.99
Round credit	Credit of 50 or more that is a multiple of 25
Exact repeat	Same folio, item, and amount within 10 minutes
Same-day repeat	Same folio, item, and amount on the same business date, after 10 minutes
Non-recurring repeat	Same folio, item, and amount within 36 hours, excluding room and breakfast
Room-night repeat	Second room charge on a folio for the same night
Clerk credit activity	Clerk's share of credits over the prior 30 days, robust z within property
Credits on folio	Number of credits already posted to the folio
Amount versus peers	Robust z of amount within property, type, and item (development medians)

2.5 Detectors

R1, policy rule set. A record is flagged if any of eight rules fires: unlinked credit; cumulative credit above the linked charge; credit more than seven days after checkout; adjustment of 50% or more of a charge; unapproved credit at or above the limit; exact repeat; same-day repeat; room-night repeat. R1 has no tunable threshold and stands for a typical exception report of the kind recommended in internal control and fraud risk guidance (COSO & ACFE, 2023; COSO, 2013).

R2, expert scorecard. Fourteen yes or no indicators carry points fixed before any results were seen: 4 points each for an exact repeat and a second room charge for the same night; 3 each for an unlinked credit, a cumulative credit above the linked charge, and an unapproved credit at or above

the limit; 2 each for a same-day repeat, a non-recurring repeat, a credit more than seven days after checkout, an adjustment of 50% or more, and an off-hours credit; and 1 each for a credit just under the limit, a round credit amount, high clerk credit activity, and two or more credits on the folio. The score is the sum, and the threshold is chosen on calibration data. The design follows the logic of points-based scorecards used in credit decisions (Siddiqi, 2006; Ustun & Rudin, 2016).

S1, peer robust z-score. Each record receives the largest of four robust z-scores: amount against its peer group, cumulative credit ratio, days after checkout, and clerk credit activity. S1 needs no labels and reflects a common audit-analytics screen based on the median and median absolute deviation (Iglewicz & Hoaglin, 1993; Leys et al., 2013).

S2, sparse logistic regression. An L1-penalized logistic regression ($C = 0.5$, balanced class weights, standardized inputs; ratios and counts log-transformed) was fitted on the January to June data with the liblinear solver (Fan et al., 2008; Friedman et al., 2010; Tibshirani, 1996). Each alert can be explained as a sum of indicator contributions, read in the usual way as odds multipliers (Hosmer et al., 2013).

S3, shallow decision tree. A classification tree with maximum depth 4, at least 50 records per leaf, and balanced class weights. Each alert follows one path of at most four questions (Breiman et al., 1984).

Reference, isolation forest. 200 trees fitted to a sample of 150,000 development records, without labels. It is included to show what an opaque, widely used method achieves on the same indicators (Liu et al., 2008, 2012).

2.6 Operating points and reviewer workload

Two thresholds per detector were fixed on the calibration months. **OP1** maximizes F1, the harmonic mean of precision and recall. **OP2** is the lowest threshold whose alert rate stays within a reviewer capacity of 30 alerts per property-week. For detectors with few distinct score values (R_1 , R_2), OP2 can coincide with OP1. Workload was translated into time with an assumed 5 minutes of triage per alert, so 30 alerts equal 2.5 hours per property-week. We report alerts per 1,000 records, alerts and hours per property-week, and alerts per true positive (the number of alerts a reviewer reads to find one real anomaly). Fixing a review budget in this way mirrors practice in card-fraud operations, where investigators can examine only a fixed number of alerts per day (Dal Pozzolo et al., 2014, 2018).

2.7 Class imbalance

For S2 and S3 we compared three training strategies: no adjustment, balanced class weights, and random undersampling of normal records to 20 per anomaly (Dal Pozzolo et al., 2015; He & Garcia, 2009; King & Zeng, 2001). Each was assessed at the default probability threshold of 0.5 and at a threshold tuned on calibration data. Separately, all detectors were evaluated with thresholds held fixed while anomaly prevalence in the evaluation period was set to 0.1%, 0.5%, and 2.0%.

2.8 Sensitivity to transaction patterns

Four pattern shifts were applied to the evaluation period only, leaving models and thresholds unchanged (Table 4).

Table 4. Evaluation scenarios.

Scenario	What changes from September to December
Baseline	Same mechanism as development; 0.5% anomalies
Low / high prevalence	Anomaly rate 0.1% or 2.0%
Group-booking surge	On 60% of days each hotel hosts a group: 6 to 14 rooms billed nightly to one master folio at the same rate
Service-recovery promotion	Allowance frequency triples (4% to 12% of folios) and 60% of allowances are 30% to 60% of the charge, with approval
15% price increase	All charge prices rise 15%, while peer medians remain those of January to June
Concealment	Anomalies shift to the subtlest variants: repeat partial refunds, threshold-hugging allowances, delayed duplicates; posted in business hours; only 40% by the two concentrated clerks

2.9 Performance measures and uncertainty

Precision is $TP / (TP + FP)$, recall is $TP / (TP + FN)$, and the false-positive rate is $FP / (FP + TN)$, where TP, FP, FN, and TN count records. PR-AUC is the average precision over the full score ranking. Because records within a property and week share guests, clerks, and events, 95% intervals were obtained by a block bootstrap that resamples whole property-weeks (300 resamples) (Efron & Tibshirani, 1993; Künsch, 1989; Politis & Romano, 1994). The full pipeline was repeated on two further datasets generated with different random seeds.

2.10 Implementation

The generator, indicators, models, and evaluation were written in Python (NumPy, pandas, scikit-learn) (Pedregosa et al., 2011), using conventions from standard statistical learning texts (Hastie et al., 2009). Every random step is seeded, and the scripts reproduce all tables and figures in this paper. The default parameter values are those stated in Sections 2.2 to 2.9.

3. Results and Discussion

3.1 Simulated ledgers

The baseline dataset contained 600,512 records: 303,836 food and beverage charges, 185,588 room charges, 104,213 other charges, 3,863 adjustments, and 3,012 refunds. The evaluation period averaged 2,639 records and 13.0 anomalies per property-week, so a perfect detector would raise 13 alerts per property-week.

3.2 Detection accuracy

Table 5 reports held-out performance at OP1. The policy rule set R1 found 88% of anomalies but generated 2,821 false alerts against 823 true ones. The sparse logistic model S2 achieved the highest F1 and PR-AUC: 84% of its alerts were real, it found 72% of anomalies, and it flagged

0.07% of normal records. The scorecard R2 was the most precise detector but missed about half of all anomalies. The decision tree S3 sat between S2 and R2. The unsupervised peer screen S1 and the isolation forest both trailed the supervised transparent models, which shows that interpretability did not cost accuracy in this setting; the opaque reference was beaten by two transparent methods.

Table 5. Held-out performance at OP1 (threshold maximizing F1 on calibration data), September to December 2025.

Detector	Precision CI)	(95% CI)	Recall (95% CI)	FPR	F1	PR- AUC	TP / FP
R1 Rule set	0.226 (0.212, 0.241)		0.878 (0.862, 0.895)	1.49 %	0.35 9	0.199	823 / 2,821
R2 Scorecard	0.945 (0.923, 0.966)		0.512 (0.479, 0.545)	0.01 %	0.66 4	0.600	480 / 28
S1 Peer robust z	0.562 (0.516, 0.604)		0.409 (0.374, 0.440)	0.16 %	0.47 3	0.377	383 / 299
S2 Sparse logistic	0.839 (0.804, 0.871)		0.722 (0.686, 0.751)	0.07 %	0.77 6	0.747	676 / 130
S3 Shallow tree	0.721 (0.682, 0.756)		0.605 (0.579, 0.636)	0.12 %	0.65 8	0.634	567 / 220
Isolation forest (ref.)	0.608 (0.575, 0.643)		0.502 (0.460, 0.535)	0.16 %	0.55 0	0.537	470 / 303

FPR, false-positive rate. Intervals from 300 block-bootstrap resamples of property-weeks. 937 anomalies among 189,988 records.

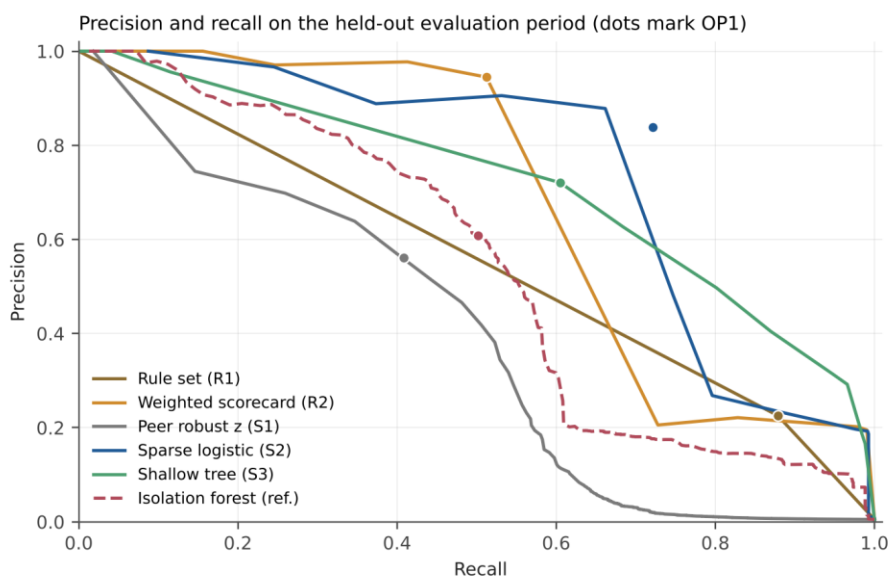


Figure 2. Precision-recall curves on the evaluation period. Dots mark the OP1 threshold chosen on calibration data. R1 has a single operating point.

3.3 Reviewer workload

Table 6 converts each detector's alerts into review time. At OP1, S2 asked for 11.2 alerts per property-week, under one hour, and reviewers read 1.19 alerts for each real anomaly. R1 asked for 4.5 times as much review for 16 percentage points more recall. Across the four hotels over a year, that difference is roughly 877 hours of triage for R1 against 194 hours for S2.

At OP2, where thresholds were relaxed to use the full 30-alert capacity on calibration data, recall rose for the statistical models (S3 to 0.802, S2 to 0.748) but precision fell by about half, and evaluation workloads of 21 to 24 alerts per property-week stayed inside capacity. Figure 3 shows the full trade-off: for any number of alerts a team can read each week, it shows the share of anomalies that would be found. S2 dominates at workloads near the true anomaly count; S3 overtakes it between roughly 17 and 60 alerts per property-week; R1 only becomes competitive above 50.

Table 6. Reviewer workload on the evaluation period at both operating points.

Detector		Op.	Prec.	Recall	FPR	Alerts / prop- week	Hours / prop- week	Alerts per TP	Hours per year (4 hotels)
R1 Rule set		fixed	0.226	0.878	1.49 %	50.6	4.22	4.43	877
R2 Scorecard		OP1 OP2	= 0.945	0.512	0.01 %	7.1	0.59	1.06	122
S1 Peer robust z		OP1	0.562	0.409	0.16 %	9.5	0.79	1.78	164
S1 Peer robust z		OP2	0.317	0.543	0.58 %	22.3	1.86	3.15	386
S2 logistic	Sparse	OP1	0.839	0.722	0.07 %	11.2	0.93	1.19	194
S2 logistic	Sparse	OP2	0.407	0.748	0.54 %	23.9	1.99	2.46	414
S3 Shallow tree		OP1	0.721	0.605	0.12 %	10.9	0.91	1.39	189
S3 Shallow tree		OP2	0.497	0.802	0.40 %	21.0	1.75	2.01	364
Isolation forest		OP1	0.608	0.502	0.16 %	10.7	0.89	1.64	186
Isolation forest		OP2	0.320	0.595	0.63 %	24.2	2.01	3.12	419

Prec., precision. Hours assume 5 minutes of triage per alert. Annual hours = hours per property-week x 4 properties x 52 weeks. R2 has no threshold between 4 and 5 points, and a threshold of 4 exceeded capacity on calibration data, so OP2 equals OP1.

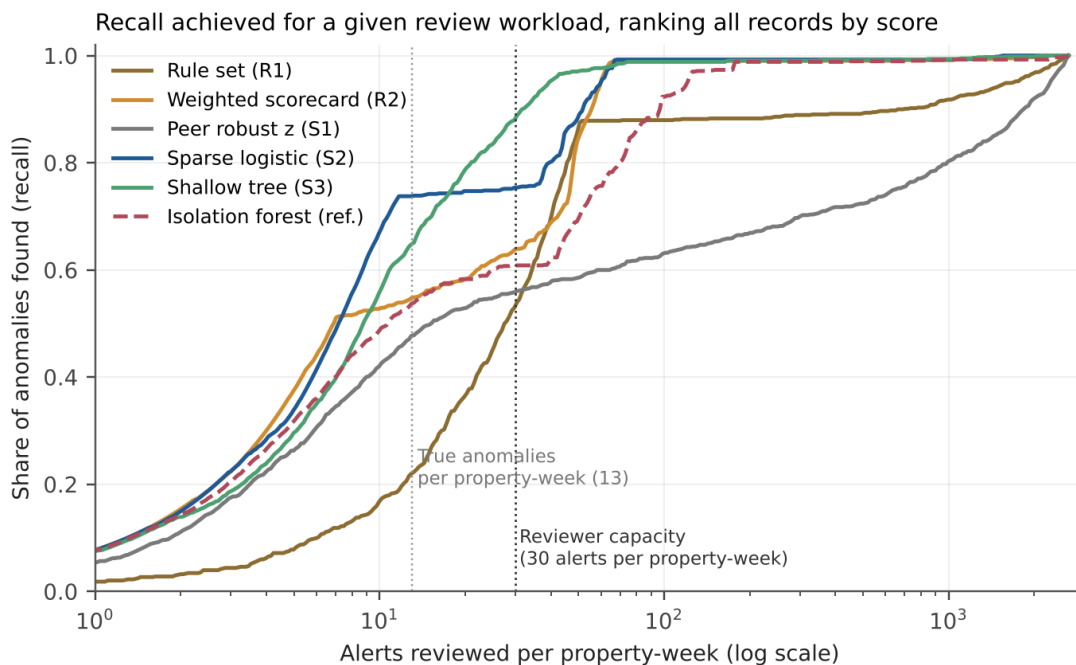


Figure 3. Recall achieved for a given weekly review workload when records are reviewed in score order. Dotted lines mark the true anomaly rate (13 per property-week) and the assumed reviewer capacity (30).

3.4 Detection by anomaly type

Table 7 breaks recall down by anomaly type. Unusual refunds were easy for supervised models (S2 recall 0.91 across refund types) and for R1, whose rules mirror their definitions. Adjustments were caught at 0.88 by S2, which was notably better than R1 on threshold hugging (0.93 against 0.72) because it combines the near-limit flag with clerk activity and posting time. Duplicates were the clear weak point: S2 found only 34% of them. Exact and delayed duplicates share every recorded field with legitimate repeat orders, so any model that keeps precision high must give these indicators modest weight. R1 caught every exact duplicate, but its duplicate rules alone produced 37.6 alerts per property-week at 7.8% precision (12.8 alerts per real duplicate). Room-night duplicates were caught by all supervised methods because legitimate folios never carry two room charges for one night. The type profile also exposes a blind spot in the tree. S3 never split on days after checkout, so it found only 14% of late refunds and 10% of repeat refunds. The unsupervised methods missed duplicates almost entirely, because a duplicated charge has a perfectly ordinary amount.

Table 7. Recall by anomaly type at OP1 on the evaluation period.

Anomaly type (n)	R1	R2	S1	S2	S3	IF
Refund: over-refund (99)	1.00	0.61	0.84	0.92	0.88	0.96
Refund: unlinked (93)	1.00	0.74	0.60	0.88	0.90	0.58
Refund: late (59)	1.00	0.29	1.00	0.81	0.14	0.95
Refund: repeat (77)	1.00	0.69	0.56	1.00	0.10	0.90
Adjustment: large share (116)	1.00	0.38	0.20	0.85	0.86	0.56
Adjustment: threshold hugging (125)	0.72	0.57	0.54	0.93	0.80	0.61
Adjustment: round unlinked (78)	1.00	0.92	0.64	0.83	0.92	0.71
Duplicate: exact (121)	1.00	0.32	0.01	0.32	0.32	0.00
Duplicate: delayed (147)	0.46	0.22	0.01	0.25	0.32	0.00
Duplicate: room night (22)	1.00	1.00	0.00	1.00	1.00	0.00
Family: refunds (328)	1.00			0.91		
Family: adjustments (319)	0.89			0.88		
Family: duplicates (290)	0.73			0.34		

3.5 Model explanations

Figure 4 shows the S2 weights as odds multipliers per standard deviation of each indicator. The penalty at $C = 0.5$ retained all 17 indicators, but four of them (posting hour, amount against peers, adjustment flag, prior credits on the folio) had multipliers between 0.93 and 1.09 and contribute little. The strongest signals were the cumulative credit ratio (3.88), an unlinked credit (2.00), and repeats of non-recurring items (1.86) or within ten minutes (1.63). Two weights are negative and need care when explaining alerts: once the cumulative ratio is known, the single-credit ratio and the refund flag act as corrections, because legitimate refunds are often large single credits that do not exceed the charge. Reviewers should read S2 explanations as a set of contributions rather than as independent rules, and a deployment could add sign constraints so that every weight agrees with audit intuition.

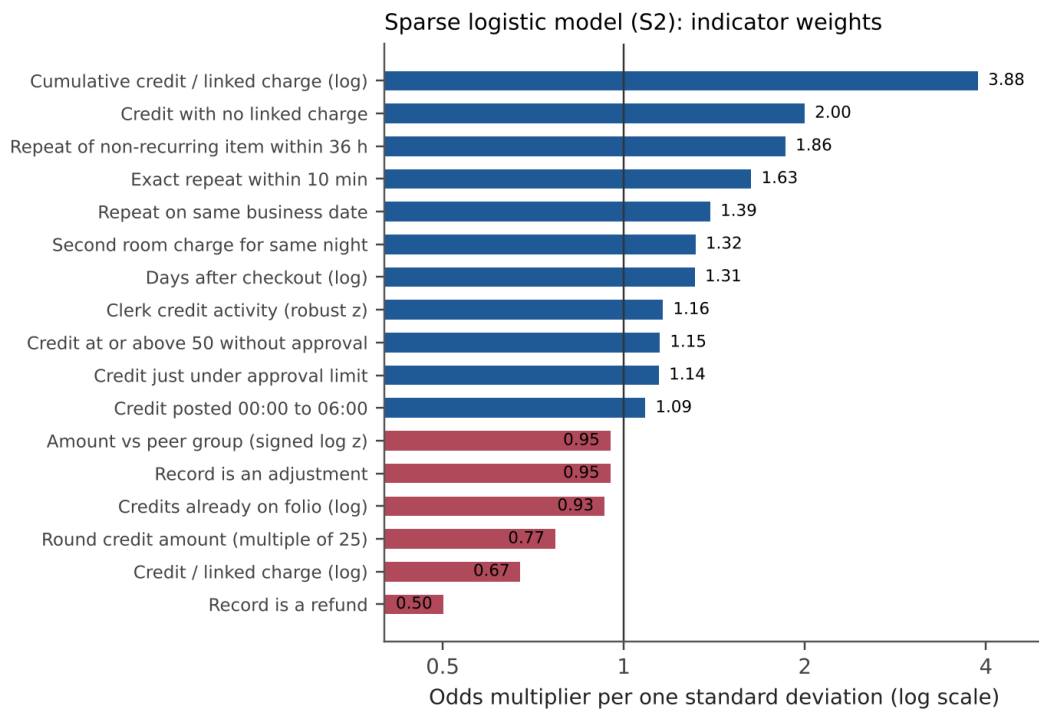


Figure 4. Weights of the sparse logistic model (S2) expressed as odds multipliers per standard deviation. Blue indicators raise the score; red indicators lower it.

The tree S3 had 12 leaves. Its first question is whether the record is a credit; non-credits are then routed by the three duplicate flags, and credits by amount against peers and cumulative credit ratio. This is easy to read but, as Table 7 showed, it ignores timing.

The scorecard's threshold sweep (Table 8) shows why R2 behaved as a high-precision filter. Every score below 5 contained large numbers of legitimate repeat bar orders (4 points for an exact repeat) and goodwill credits; above 5, nearly every alert was real. A reviewer can see in one line why an alert scored 5: for example, an unapproved credit (3) posted off-hours (2).

Table 8. Scorecard (R2) operating points on the evaluation period.

Minimum score	1	2	3	4	5	6	7	8
Precision	0.195	0.202	0.221	0.206	0.945	0.977	0.971	1.000
Recall	0.990	0.979	0.828	0.728	0.512	0.413	0.247	0.156
Alerts per property-week	66.0	63.1	48.7	46.1	7.1	5.5	3.3	2.0

3.6 Class imbalance

Table 9 shows that handling imbalance changed thresholds more than it changed ranking. At the default threshold of 0.5, balanced class weights turned S2 into a near-exhaustive screen (recall 0.977) with 65 alerts per property-week at 19.5% precision, because weighting inflates predicted probabilities. Once the threshold was tuned on calibration data, all three strategies settled at 10 to 12 alerts per property-week. Contrary to common practice, the unadjusted model ranked anomalies best (PR-AUC 0.845 against 0.747 with weighting) and reached precision 0.907 with recall 0.685

after tuning. Undersampling was close to the unadjusted model and trains on far fewer records. The pre-specified weighted model, used as S2 elsewhere in this paper, is therefore a conservative choice.

Table 9. Class-imbalance strategies for the statistical models, evaluation period.

Model and strategy	PR-AUC	Precision at 0.5	Recall at 0.5	Alerts / pw at 0.5	Precision tuned	Recall tuned	Alerts / pw tuned
S2, no adjustment	0.845	0.888	0.705	10.3	0.907	0.685	9.8
S2, class weighting	0.747	0.195	0.977	65.1	0.839	0.721	11.2
S2, undersampling 1:20	0.834	0.454	0.870	25.0	0.818	0.755	12.0
S3, no adjustment	0.665	0.873	0.674	10.1	0.873	0.674	10.1
S3, class weighting	0.634	0.188	0.984	68.1	0.720	0.605	10.9
S3, undersampling 1:20	0.717	0.425	0.859	26.3	0.774	0.689	11.6

Tuned thresholds maximize F1 on July to August data. pw, property-week. The S2 class-weighting row comes from a separate refit and matches Table 5 to within 0.001.

Prevalence had a larger effect than training strategy. With thresholds held fixed, recall barely moved as the anomaly rate changed, but precision tracked prevalence closely (Figure 5). At 0.1% prevalence, S2 precision fell to 0.57 and R1 precision to 0.06, meaning about 16 alerts for every real one; the scorecard kept precision at 0.86 but on only 1.8 alerts per property-week. At 2.0% prevalence, S2 precision rose to 0.92 and its workload to 41 alerts per property-week, above the assumed capacity.

3.7 Sensitivity to transaction patterns

Figure 5 summarizes recall, precision, and workload for every detector and scenario at OP1. Three failures stand out.

- **Price change broke the peer screen.** A 15% price increase pushed most charges beyond the January to June peer medians. S1 then flagged 14% of all records (375 alerts per property-week, precision 0.02). Detectors that use ratios and flags rather than absolute amounts were unaffected; S2 kept precision 0.81 and recall 0.71.
- **Group billing broke duplicate keys.** Master folios that carry several rooms at the same rate on the same night look exactly like room-night and exact duplicates when the key is folio, item, and amount. Every detector that uses those flags collapsed: S2 precision fell to 0.09 with 98 alerts per property-week. Adding the room number to the duplicate key restored S2 to precision 0.85 at 9.9 alerts, and R2 to 0.93. R1 returned to 43.5 alerts per property-week, close to its usual level.

- **A legitimate promotion inflated adjustment alert.** When approved allowances of 30% to 60% became common, the tree S3 (which leans on amount against peers) doubled its alerts to 19.7 per property-week and its precision fell to 0.37. S2 was more robust (precision 0.75 at 12.4 alerts), because approval status offsets the large ratio.

Under concealment, where anomalies used only their subtlest forms, R1 recall fell from 0.88 to 0.75, S3 from 0.61 to 0.40, and R2 from 0.51 to 0.42, while S2 held at 0.71 with unchanged precision (0.85). The isolation forest was stable across most scenarios, but at a consistently lower level than S2.

Sensitivity of each method to evaluation scenarios (thresholds fixed from calibration, OP1)

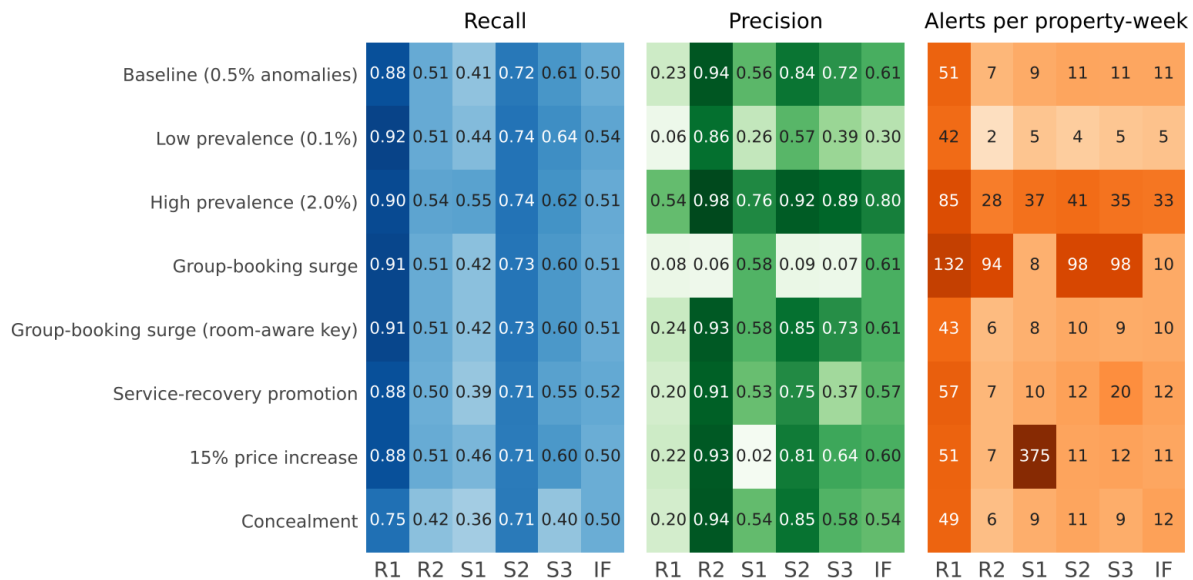


Figure 5. Recall, precision, and alerts per property-week for each detector (columns) under each evaluation scenario (rows), with thresholds fixed from calibration (OP1). IF, isolation forest.

3.8 Replication

Repeating the full pipeline on two more datasets gave the same ordering (Table 10). S2 precision ranged from 0.836 to 0.861 and recall from 0.666 to 0.738 across the three datasets. The tree was the least stable method, with recall from 0.605 to 0.741 and workload from 10.9 to 17.0 alerts per property-week, a familiar property of single trees whose top splits change with small changes in data.

Table 10. Performance at OP1 across three independently generated datasets: mean (minimum, maximum).

Detector	Precision	Recall	PR-AUC	Alerts / property-week
R1 Rule set	0.225 (0.225, 0.226)	0.885 (0.878, 0.888)	0.200	51.8 (50.6, 52.9)
R2 Scorecard	0.934 (0.923, 0.945)	0.512 (0.504, 0.520)	0.594	7.2 (7.1, 7.5)
S1 Peer robust z	0.547 (0.535, 0.562)	0.465 (0.409, 0.519)	0.411	11.3 (9.5, 13.0)
S2 Sparse logistic	0.845 (0.836, 0.861)	0.708 (0.666, 0.738)	0.748	11.1 (10.2, 11.9)
S3 Shallow tree	0.680 (0.575, 0.743)	0.655 (0.605, 0.741)	0.618	13.0 (10.9, 17.0)
Isolation forest	0.660 (0.608, 0.687)	0.496 (0.487, 0.502)	0.547	10.0 (9.4, 10.7)

3.9 Two-tier review queue

The results suggest pairing the scorecard's precision with the logistic model's coverage. In the recommended design, **Tier 1** contains records with a scorecard score of 5 or more, each delivered with the list of indicators that earned its points. **Tier 2** contains the remaining S2 alerts at OP1, each delivered with its three largest indicator contributions. Both thresholds were fixed before evaluation, so Table 11 is a held-out result for the combined model.

Table 11. Evaluated two-tier model and its components, evaluation period.

Queue	Precision	Recall	FPR	Alerts prop-week	Hours prop-week	Alerts per TP
Tier 1: scorecard score 5 or more	0.945	0.512	0.01%	7.1	0.59	1.06
Tier 2: added S2 alerts (OP1)	0.658	0.226	0.06%	4.5	0.37	1.52
Combined two-tier model	0.834	0.739	0.07%	11.5	0.96	1.20
For comparison: R1 refund and adjustment rules only	0.655	0.653	0.17%	13.0	1.08	1.53
For comparison: R1 duplicate rules only	0.078	0.225	1.32%	37.6	3.14	12.84

Recall of Tier 2 is the additional share of all anomalies found beyond Tier 1.

The combined model found 74% of anomalies with 83% of alerts real, at 11.5 alerts and just under one hour per property-week, about 200 hours a year across the four hotels. It matched S2 alone on

workload while ordering the work so that the most certain cases come first. Its remaining gap is duplicates, which Section 3.11 discusses.

3.10 Principal findings

In this simulation, transparent methods identified unusual refunds and adjustments well and at modest workloads. A sparse logistic model built on 17 checkable indicators found about 72% of all anomalies, and about 90% of refund and adjustment anomalies, while asking a reviewer to read roughly 11 records per hotel per week. An expert scorecard identified half of all anomalies with almost no false alerts. A two-tier queue that combines them reached 74% recall at 83% precision for under an hour of review per property-week. A conventional exception report achieved higher recall but at four to five times the workload, mostly through duplicate rules that flag legitimate repeat orders. How well transparent detection works depends on the anomaly family. For refunds and adjustments, the information needed to separate legitimate from irregular activity is present in the ledger: link to an original charge, cumulative amount, timing, approval, and who posted it. For duplicates, the ledger fields used here are not enough, and better models do not fix missing information. The opaque isolation forest ranked below two transparent supervised methods on every headline measure. This supports the view that, for structured records with meaningful features, interpretable models can perform as well as complex ones (Breiman, 2001; Caruana et al., 2015; Rudin, 2019). This is also what audit guidance for AI-supported financial systems asks for: evidence that can be validated and human oversight of automated judgments (Elebe & Anunagba, 2025; Fobellah, 2025b; Ominyi & Anichukwueze, 2023). The benefit of transparency here is practical: every alert in the recommended model arrives with reasons a reviewer can check against the folio in a few minutes, and the same reasons can be shared with the employee involved.

3.11 Duplicate entries and data keys

Two results point the same way. Duplicate detection was poor when legitimate repeats shared all key fields, and it collapsed when group billing created more such repeats; adding the room number to the key fixed the second problem completely. In practice, the point-of-sale check number, the terminal, and the original posting reference would distinguish a reposted check from a second round of drinks. Hotels considering duplicate analytics should first confirm that these identifiers flow into the ledger extract, which is a data integration and governance task as much as an analytic one (Aliliele et al., 2025b; Badmus et al., 2019; Mbonu et al., 2021a). Without them, duplicate alerts should go to a separate low-priority queue so that they do not crowd out credit alerts.

3.12 Thresholds and changing transaction patterns

The imbalance experiment showed that class weighting mostly changes where a threshold falls, and that threshold choice, not training strategy, determines workload. The prevalence experiment showed that a threshold that is right at 0.5% anomalies gives far more false alerts per real one at 0.1%. Thresholds should therefore be owned by the finance team, documented, and reviewed on a schedule, in the same way risk-based internal audit treats other control parameters (Akomolafe & Agu, 2018; Akomolafe et al., 2025; Ominyi & Anichukwueze, 2024), and in line with guidance on recalibration and drift monitoring (Gama et al., 2014; Van Calster et al., 2019), using recent labeled outcomes from the review queue. A simple control is to track weekly alerts and confirmed findings and to reset the threshold when either drifts beyond an agreed band.

Each failure in Section 3.7 traces to a known business event: a rate increase, a group contract, a service-recovery campaign. These events are announced in advance, so detectors can be stress-

tested before they occur, as this study did. Ratio-based indicators survived the price change while absolute peer baselines did not, which argues for rebasing peer statistics after any price change and for preferring relative indicators where possible. This is the dataset-shift problem described in the machine learning literature, here made concrete with hotel events (Gama et al., 2014; Moreno-Torres et al., 2012).

3.13 Implications for practice and research

Earlier work in this line of inquiry assessed revenue-integrity risk conceptually: which points in the hotel revenue cycle are exposed, which controls are expected, and which red flags auditors should watch. That work sits within a wider body of conceptual models for technology-enabled risk assessment, audit analytics, and transaction anomaly detection (Abetoh & Atakpa, 2024; Akomolafe & Agu, 2018; Atakpa & Fobellah, 2023; Ominyi et al., 2024), most of which propose architectures without measuring how they perform. This study moves that assessment to testable transaction analytics. Each conceptual red flag became a computable indicator; each control expectation became a rule or a weight; and each claim about detection became a measured precision, recall, false-positive rate, and workload on data that the methods had not seen. The scenario generator also turns qualitative concerns, such as group billing or promotions, into tests that can be repeated whenever a property changes how it operates. The natural next step is a pilot on a hotel's own ledger extract, using the two-tier model as a starting point and the review outcomes as labels.

3.14 Limitations and future research

The data are synthetic. The anomaly definitions, the indicators, and the policy rules were written by the same authors, which favors the rule set and the supervised models; R1's perfect recall on most refund types reflects this. Absolute numbers should not be read as forecasts for a real hotel. The relative findings (precision of scorecards, the duplicate problem, threshold sensitivity, the failure modes under pattern shift) are more likely to transfer, because they arise from the structure of ledger data rather than from tuning. This is the usual standard for inference from simulation (Morris et al., 2019; Sargent, 2013). The supervised models assume labeled anomalies exist for training. Real hotels have few confirmed cases, and labels from past reviews are biased toward what earlier rules flagged. In practice the scorecard and peer screen would be used first, and a logistic model would be fitted once enough reviewed alerts had accumulated, in the human-in-the-loop manner described for other review workflows (Ladapo et al., 2022, 2024). Card-fraud studies show that such feedback loops also bias future labels, which must be managed (Dal Pozzolo et al., 2018). The workload model assumes 5 minutes per alert for every alert. Real triage time varies by alert type, and confirmed anomalies need further investigation that is not counted here. The capacity of 30 alerts per property-week is an assumption that each hotel should replace with its own figure. Only one penalty setting and one tree depth were examined, and the scorecard points were fixed by judgment rather than optimized. Other transparent models, such as generalized additive models or rule lists, might perform better (Caruana et al., 2015; Ustun & Rudin, 2016). Finally, the study treats each record independently, whereas card-fraud work shows that aggregating transactions over time and across entities can add signal; schemes that spread across many small transactions or several clerks may need folio-level or user-level aggregation.

4. Conclusion

Transparent methods can find most unusual refunds and adjustments in hotel revenue records with review burdens a finance team can sustain. In a controlled simulation withheld-out evaluation, a sparse logistic model on 17 checkable indicators reached 84% precision and 72% recall at about 11 alerts and under one hour of review per property-week, and a two-tier queue combining it with an expert scorecard reached 83% precision and 74% recall at similar cost. Exception-style rules found more anomalies but quadrupled the workload. Duplicate entries remain hard to separate from legitimate repeats without point-of-sale identifiers. Thresholds must be managed as prevalence changes, peer baselines must be refreshed after price changes, and duplicate keys must reflect how a property bills group. With those safeguards, explainable detection can support routine, testable revenue assurance in hotel operations.

References

- Abdallah, A., Maarof, M. A., & Zainal, A. (2016). Fraud detection system: A survey. *Journal of Network and Computer Applications*, 68, 90 to 113.
- Abetoh, N. F., & Atakpa, M. I. (2024). Audit analytics in healthcare financial oversight: Leveraging data science to strengthen accountability in multilateral grant ecosystems. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 10(6), 2710 to 2747. <https://doi.org/10.32628/CSEIT2410791>
- Adelanwa, A., Basnet, A., & Anene, U. N. (2023). Predictive analytics models for financial risk detection and fraud prevention in public systems. *International Journal of Advanced Multidisciplinary Research and Studies*, 3(6). <https://doi.org/10.62225/2583049X.2023.3.6.5969>
- Adesuyi, M. O., Akomolafe, O., Olaogun, B. O., Ndukwe, V. U., & Sakyi, J. K. (2024). AI-driven risk scoring model for global cross-border trade payment transactions. *International Journal of Advanced Multidisciplinary Research and Studies*, 4(1), 1569 to 1581. <https://doi.org/10.62225/2583049X.2024.4.1.5281>
- Adesuyi, M. O., Akomolafe, O., Olaogun, B. O., Ndukwe, V. U., & Sakyi, J. K. (2025). AI-enabled fraud detection ecosystem model for securing international payment channels. *International Journal of Advanced Multidisciplinary Research and Studies*, 5(6), 866 to 882. <https://doi.org/10.62225/2583049X.2025.5.6.5284>
- Adesuyi, M. O., Kalu, A., & Walawalkar, G. (2023). Data-led cost governance in technology-intensive enterprises. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 9(3), 877 to 896. <https://doi.org/10.32628/IJSRCSEIT>
- Adesuyi, M. O., Walawalkar, G., & Kalu, A. (2022). Predictive budgeting models using operational and market signals. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 8(5), 818 to 837. <https://doi.org/10.32628/IJSRCSEIT>
- Agboada, P., Ezeako, R., Okoli, I. E., & Uzundu, C. S. (2018). Digital transformation in operations management: A review of advances in automation, analytics, and process redesign. *International Journal of Engineering Technology Research & Management*, 2(11), 109 to 134.
- Agboada, P., Ezeako, R., Okoli, I. E., & Uzundu, C. S. (2019). Standard operating procedures as change infrastructure: A conceptual framework for sustaining operational improvement. *Iconic Research and Engineering Journals*, 3(6), 706 to 730. <https://doi.org/10.64388/IREV316-1723313>
- Agu, M. U., Akomolafe, O., & Bello, A. (2023). A comparative review of SOX compliance frameworks in cross-border financial auditing. *International Journal of Advanced Multidisciplinary Research and Studies*, 3(6), 2297 to 2306. <https://doi.org/10.62225/2583049X.2023.3.6.5362>
- Agu, M. U., Akomolafe, O., & Bello, A. (2025). Advances in predictive financial risk assessment using Python-based forecasting systems. *International Journal of Computer Science and Mathematical Theory*, 11(11), 131 to 148. <https://doi.org/10.56201/ijcsmt.vol.11.no11.2025.pg131.148>
- Akhigbe, R., Falemi, A., & Akin-Oluyomi, O. T. (2023). A conceptual model for improving customer experience using workforce behavior and real-time operational data.

- International Journal of Multidisciplinary Research and Growth Evaluation*, 4(6), 1322 to 1338. <https://doi.org/10.54660/ijmrge.2023.4.6.1322-1338>
- Akin-Oluyomi, O. T., & Akhigbe, R. (2022). A supply chain governance model for enhancing compliance and operational quality across retail networks. *International Journal of Multidisciplinary Research and Growth Evaluation*, 3(6), 860 to 878. <https://doi.org/10.54660/ijmrge.2022.3.6.860-878>
- Akin-Oluyomi, O. T., & Akhigbe, R. (2023). An advanced framework for improving multi-site operational efficiency using data-driven performance indicators. *International Journal of Advanced Multidisciplinary Research and Studies*, 3(1), 1763 to 1781. <https://doi.org/10.62225/2583049X.2023.3.1.5343>
- Akinleye, O. K., & Adeyoyin, O. (2021). Process automation framework for enhancing procurement efficiency and transparency. *Shodhshauryam, International Scientific Refereed Research Journal*, 4(4), 356 to 387.
- Akomolafe, O., & Agu, M. U. (2018). A conceptual model for enhancing internal audit quality through technology-enabled risk assessment frameworks. *Iconic Research and Engineering Journals*, 1(9), 458 to 475.
- Akomolafe, O., & Agu, M. U. (2019a). A conceptual framework for developing risk-based internal control models in the insurance and banking sectors. *Iconic Research and Engineering Journals*, 2(8), 335 to 354.
- Akomolafe, O., & Agu, M. U. (2019b). A review of data-driven risk evaluation models for emerging market financial institutions. *Iconic Research and Engineering Journals*, 3(6), 433 to 448.
- Akomolafe, O., & Agu, M. U. (2019c). Advances in financial resilience through integrated governance and compliance strategies. *Iconic Research and Engineering Journals*, 2(10), 607 to 620.
- Akomolafe, O., Agu, M. U., & Bello, A. (2023a). A conceptual model for implementing risk-based auditing in strategic financial management. *International Journal of Advanced Multidisciplinary Research and Studies*, 3(6), 2274 to 2286. <https://doi.org/10.62225/2583049X.2023.3.6.5359>
- Akomolafe, O., Agu, M. U., & Bello, A. (2023b). A quantitative conceptual model for assessing compliance risk in emerging economies. *International Journal of Advanced Multidisciplinary Research and Studies*, 3(6), 2287 to 2296. <https://doi.org/10.62225/2583049X.2023.3.6.5360>
- Akomolafe, O., Agu, M. U., & Bello, A. (2025). A conceptual model for advancing risk governance through data-driven compliance analytics in financial institutions. *Journal of Accounting and Financial Management*, 11(11), 211 to 228. <https://doi.org/10.56201/jafm.vol.11.no11.2025.pg211.228>
- Alieme, J. I., Sule, A., & Obukadata, V. O. (2022). Standardization of reporting and decision workflows: A framework for improving efficiency, data quality, and management visibility. *International Journal of Multidisciplinary Research and Growth Evaluation*, 3(6), 1190 to 1218. <https://doi.org/10.54660/IJMRGE.2022.3.6.1190-1218>
- Aliliele, C., Eboh, E. E., & Oluwo, K. (2024b). Advances in variance analytics for cost control in manufacturing and industrial enterprises. *International Journal of Advanced Multidisciplinary Research and Studies*, 4(6), 3163 to 3185. <https://doi.org/10.62225/2583049X.2024.4.6.6012>

- Aliliele, C., Mbonu, I. S., & Iwuanyanwu, U. (2023). A review of API governance and risk prioritization frameworks in modern financial institutions. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 9(10), 395 to 433. <https://doi.org/10.32628/CSEIT2361072>
- Aliliele, C., Mbonu, I. S., & Iwuanyanwu, U. (2024a). A conceptual framework for enterprise data sensitivity classification and regulatory traceability mechanisms. *International Journal of Advanced Multidisciplinary Research and Studies*, 4(6), 3103 to 3124. <https://doi.org/10.62225/2583049X.2024.4.6.5991>
- Aliliele, C., Mbonu, I. S., Uzoka, E., & Iwuanyanwu, U. (2025a). A review of AI assisted continuous auditing systems in technology risk and cybersecurity oversight. *Gyanshauryam, International Scientific Refereed Research Journal*, 8(4), 210 to 250. <https://doi.org/10.32628/GISRRJ258369>
- Aliliele, C., Mbonu, I. S., Uzoka, E., & Iwuanyanwu, U. (2025b). Advances in data lakehouse governance architectures for enterprise data loss prevention and compliance assurance. *Shodhshauryam, International Scientific Refereed Research Journal*, 8(4), 171 to 213. <https://doi.org/10.32628/SHISRRJ258474>
- Alles, M. G., Kogan, A., & Vasarhelyi, M. A. (2008). Putting continuous auditing theory into practice: Lessons from two pilot implementations. *Journal of Information Systems*, 22(2), 195 to 214.
- Ambali, K. B., Eyetsemitan, R. A., Oyeleye, A. O., & Fadayomi, O. (2021). Lean Six Sigma for small enterprises: A systematic review and Lite-DMAIC adaptation framework for resource-constrained organizations. *Iconic Research and Engineering Journals*, 5(5), 562 to 583. <https://doi.org/10.64388/IREV5I5-1716957>
- Anichukwueze, C. C., Osuji, V. C., & Oguntegbe, E. E. (2021). Blockchain-based architectures for tamper-proof regulatory recordkeeping and real-time audit readiness. *International Journal of Multidisciplinary Research and Growth Evaluation*, 2(6), 485 to 504.
- Annan, A. O. (2024). Algorithmic accountability and trade secret protection in artificial intelligence. *Shodhshauryam, International Scientific Refereed Research Journal*, 7(5), 315 to 347.
- Annan, A. O. (2025). Automated decision-making and anti-discrimination compliance under U.S. law. *International Journal of Advanced Multidisciplinary Research and Studies*, 5(2), 2522 to 2540.
- Antonio, N., de Almeida, A., & Nunes, L. (2017). Predicting hotel booking cancellation to decrease uncertainty and increase revenue. *Tourism & Management Studies*, 13(2), 25 to 39.
- Antonio, N., de Almeida, A., & Nunes, L. (2019). Hotel booking demand datasets. *Data in Brief*, 22, 41 to 49.
- 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 to 27.
- Association of Certified Fraud Examiners. (2024). *Occupational fraud 2024: A report to the nations*. ACFE.
- Atakpa, M. I. (2021). A review of predictive analytics models for anti-money laundering detection in commercial banking systems. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 7(2), 778 to 804. <https://doi.org/10.32628/CSEIT2064813>

- Atakpa, M. I., & Abetoh, N. F. (2022). A systematic review of machine learning advances in financial fraud detection for banking systems. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 8(2), 771 to 801. <https://doi.org/10.32628/CSEIT23906220>
- Atakpa, M. I., & Abolaji, T. O. (2022). A privacy-preserving data architecture model for regulated industry analytics under GDPR and HIPAA compliance. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 8(3), 781 to 808. <https://doi.org/10.32628/CSEIT22558>
- Atakpa, M. I., & Fobellah, A. N. (2023). Anomaly detection in financial time-series data: A conceptual model for healthcare and banking applications. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 9(2), 967 to 997. <https://doi.org/10.32628/CSEIT2342441>
- Atakpa, M. I., Abetoh, N. F., & Akeju, B. (2024). Advances in artificial intelligence for healthcare payment fraud detection: A review of NHS applications. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 10(4), 1161 to 1194. <https://doi.org/10.32628/CSEIT26123240>
- Babatope, O. M., Akokodaripon, D. A., Akinleye, O. K., & Okoruwa, P. O. (2023). The role of data visualization in enhancing strategic procurement decision-making effectiveness. *International Journal of Advanced Multidisciplinary Research and Studies*, 3(6), 2205 to 2212.
- Badmus, O., Dosunmu, A. A., & Ozowara, D. E. (2019). A conceptual model for ETL design and data integration in Salesforce-centric enterprise architectures. *Iconic Research and Engineering Journals*, 3(5).
- Baesens, B., Van Vlasselaer, V., & Verbeke, W. (2015). *Fraud analytics using descriptive, predictive, and social network techniques: A guide to data science for fraud detection*. Wiley.
- 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 to 235.
- Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., García, S., Gil-López, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2020). Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, 58, 82 to 115.
- Bello, A. A., Oduro, D. A., Manu, E. O., Bello, A. D., Leo, A. O., Ukatu, C. E., & Okika, N. (2025c). Enhancing Know Your Customer (KYC) and Anti-Money Laundering (AML) compliance using blockchain: A business analysis approach. *Iconic Research and Engineering Journals*, 8(9), 297 to 305.
- Bello, A. D., Elebe, O., Hammed, N. I., Okoruwa, P. O., Fadayomi, O., & Omoegun, G. O. (2025b). An advanced regulatory technology framework for improving financial transparency and fraud reporting accuracy. *International Journal of Advanced Multidisciplinary Research and Studies*, 5(6).
- Bello, A. D., Oguntola, O. B., Ajibade, A. T., Akindolani, A., Ayoola, O., & Bello, A. M. (2025a). AI-driven fraud detection in UK digital payment systems: Challenges and solutions. *Iconic Research and Engineering Journals*, 9(1), 1329 to 1343.
- Bhattacharyya, S., Jha, S., Tharakunnel, K., & Westland, J. C. (2011). Data mining for credit card fraud: A comparative study. *Decision Support Systems*, 50(3), 602 to 613.

- Bola-Sadipe, D., Aliliele, C., & Eboh, E. E. (2019). Conceptual model for strategic accounting systems strengthening financial transparency and regulatory compliance. *IRE Journals*, 3(6), 535 to 564. <https://doi.org/10.64388/IREV3I6-1715327>
- Bola-Sadipe, D., Aliliele, C., & Odihi, F. (2023). Conceptual model for automated financial consolidation and transparency in multinational corporations. *Gyanshauryam, International Scientific Refereed Research Journal*, 6(6), 414 to 455. <https://doi.org/10.32628/GISRRJ236645>
- Bolton, R. J., & Hand, D. J. (2002). Statistical fraud detection: A review. *Statistical Science*, 17(3), 235 to 255.
- Breiman, L. (2001). Statistical modeling: The two cultures. *Statistical Science*, 16(3), 199 to 231.
- Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and regression trees*. Wadsworth.
- Burton, A., Altman, D. G., Royston, P., & Holder, R. L. (2006). The design of simulation studies in medical statistics. *Statistics in Medicine*, 25(24), 4279 to 4292.
- Bussmann, N., Giudici, P., Marinelli, D., & Papenbrock, J. (2021). Explainable machine learning in credit risk management. *Computational Economics*, 57(1), 203 to 216.
- Campos, G. O., Zimek, A., Sander, J., Campello, R. J. G. B., Micenková, B., Schubert, E., Assent, I., & Houle, M. E. (2016). On the evaluation of unsupervised outlier detection: Measures, datasets, and an empirical study. *Data Mining and Knowledge Discovery*, 30(4), 891 to 927.
- Carcillo, F., Le Borgne, Y.-A., Caelen, O., Kessaci, Y., Oblé, F., & Bontempi, G. (2021). Combining unsupervised and supervised learning in credit card fraud detection. *Information Sciences*, 557, 317 to 331.
- Caruana, R., Lou, Y., Gehrke, J., Koch, P., Sturm, M., & Elhadad, N. (2015). Intelligible models for healthcare: Predicting pneumonia risk and hospital 30-day readmission. In *Proceedings of the 21st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 1721 to 1730).
- Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. *ACM Computing Surveys*, 41(3), Article 15.
- Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, 16, 321 to 357.
- Committee of Sponsoring Organizations of the Treadway Commission, & Association of Certified Fraud Examiners. (2023). *Fraud risk management guide* (2nd ed.). COSO.
- Committee of Sponsoring Organizations of the Treadway Commission. (2013). *Internal control: Integrated framework*. COSO.
- Cressey, D. R. (1953). *Other people's money: A study in the social psychology of embezzlement*. Free Press.
- Cross, R. G., Higbie, J. A., & Cross, D. Q. (2009). Revenue management's renaissance: A rebirth of the art and science of profitable revenue generation. *Cornell Hospitality Quarterly*, 50(1), 56 to 81.
- Dal Pozzolo, A., Boracchi, G., Caelen, O., Alippi, C., & Bontempi, G. (2018). Credit card fraud detection: A realistic modeling and a novel learning strategy. *IEEE Transactions on Neural Networks and Learning Systems*, 29(8), 3784 to 3797.

- Dal Pozzolo, A., Caelen, O., Johnson, R. A., & Bontempi, G. (2015). Calibrating probability with undersampling for unbalanced classification. In *2015 IEEE Symposium Series on Computational Intelligence* (pp. 159 to 166). IEEE.
- Dal Pozzolo, A., Caelen, O., Le Borgne, Y.-A., Waterschoot, S., & Bontempi, G. (2014). Learned lessons in credit card fraud detection from a practitioner perspective. *Expert Systems with Applications*, 41(10), 4915 to 4928.
- Davis, J., & Goadrich, M. (2006). The relationship between precision-recall and ROC curves. In *Proceedings of the 23rd International Conference on Machine Learning* (pp. 233 to 240).
- Debreceeny, R. S., & Gray, G. L. (2010). Data mining journal entries for fraud detection: An exploratory study. *International Journal of Accounting Information Systems*, 11(3), 157 to 181.
- Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General*, 144(1), 114 to 126.
- Doshi-Velez, F., & Kim, B. (2017). *Towards a rigorous science of interpretable machine learning* (arXiv:1702.08608). arXiv.
- Ebhøjie, O., Dogbatsey, E. A., & Oyeleye, A. O. (2023). Audit liaison, corrective action planning, and control compliance in multinational organisations: A systematic literature review. *International Journal of Advanced Multidisciplinary Research and Studies*, 3(6), 2839 to 2850. <https://doi.org/10.62225/2583049X.2023.3.6.6200>
- Eboh, E. E., & Aliliele, C. (2023). Portfolio risk optimization models for venture capital investment strategy and fraud exposure analysis. *Iconic Research and Engineering Journals*, 7(4), 759 to 792. <https://doi.org/10.64388/IREV714-1715305>
- Eboh, E. E., & Aliliele, C. (2024). AI-driven data analytics framework for risk assessment and detection of venture capital and private equity investment fraud in U.S. capital markets. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 10(6), 2624 to 2665. <https://doi.org/10.32628/CSEIT2410787>
- Eboh, E. E., & Aliliele, C. (2025). Predictive analytics systems for investment risk monitoring in SME FinTech companies and cross-border capital markets. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 11(2), 3969 to 4010. <https://doi.org/10.32628/CSEIT25113398>
- Eboh, E. E., Aliliele, C., & Odihi, F. (2025). Advances in financial governance and accountability systems in small and medium enterprises. *International Journal of Advanced Multidisciplinary Research and Studies*, 5(6), 2223 to 2245. <https://doi.org/10.62225/2583049X.2025.5.6.6013>
- Edivri, J., & Oteri, O. (2024). A conceptual KPI-driven decision and optimization framework for IT service delivery, portfolio performance, and adoption. *Gyanshauryam, International Scientific Refereed Research Journal*, 7(1), 167 to 187. <https://doi.org/10.32628/GISRRJ246643>
- Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. Chapman & Hall.
- Ekechi, N. V., Anunagba, C. O., & Ozowara, D. E. (2022). Advances in financial forensics techniques for detecting cyber-enabled fraud in telemedicine services. *Shodhshauryam, International Scientific Refereed Research Journal*, 5(1), 428 to 450.
- Elebe, O., & Anunagba, C. O. (2025). Responsible artificial intelligence in IT auditing through evidence validation, human oversight, and regulatory compliance. *International Journal of Technology, Health and Sustainability*, 1(2), 310 to 328.

- Elkan, C. (2001). The foundations of cost-sensitive learning. In *Proceedings of the 17th International Joint Conference on Artificial Intelligence* (pp. 973 to 978).
- Enshara, S. I., & Bastian, I. (2018). Rasionalisasi kecurangan penerimaan kas (studi kasus pada Hotel Novotel Yogyakarta) [Rationalization of cash receipt fraud: A case study at Novotel Hotel Yogyakarta]. *ABIS: Accounting and Business Information Systems Journal*, 6(3). <https://doi.org/10.22146/abis.v6i3.59089>
- Ezeako, R., Ibeh, A. V., Alugwe, S., & Awulu, I. (2024). Performance dashboards and data-driven decision-making in operations management: A conceptual framework for leadership visibility. *International Journal of Advanced Multidisciplinary Research and Studies*, 4(6), 3520 to 3538. <https://doi.org/10.62225/2583049X.2024.4.6.6930>
- Fadayomi, O., Bello, A. D., Elebe, O., Hammed, N. I., & Omoegun, G. O. (2024). An adaptive fraud risk scoring model for real-time transaction monitoring at scale. *IIARD International Journal of Banking and Finance Research*, 10(10), 212 to 230.
- Fan, R.-E., Chang, K.-W., Hsieh, C.-J., Wang, X.-R., & Lin, C.-J. (2008). LIBLINEAR: A library for large linear classification. *Journal of Machine Learning Research*, 9, 1871 to 1874.
- Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8), 861 to 874.
- Filani, O. M., Nnabueze, S. B., Sakyi, J. K., & Okojie, J. S. (2023). Scenario-based financial modelling for enhancing strategic decision-making and organizational long-term planning. *Journal of Frontiers in Multidisciplinary Research*, 4(2), 251 to 265. <https://doi.org/10.54660/JFMR.2023.4.2.251-265>
- Fobellah, A. N. (2025a). Cognitive biases in financial decision-making: Implications for audit and risk management in large corporations: A conceptual review. *International Journal of Science and Research Archive*, 16(2), 17 to 22. <https://doi.org/10.30574/ijrsra.2025.16.2.2273>
- Fobellah, A. N. (2025b). Navigating digital transformation: Auditing artificial intelligence-powered financial systems: A conceptual review. *International Journal of Science and Research Archive*, 16(2), 23 to 28. <https://doi.org/10.30574/ijrsra.2025.16.2.2274>
- Fobellah, A. N. (2025c). Unintended consequences of financial regulations: A study of corporate compliance burdens: A conceptual review. *International Journal of Science and Research Archive*, 16(2), 29 to 34. <https://doi.org/10.30574/ijrsra.2025.16.2.2275>
- Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, 33(1), 1 to 22.
- Gama, J., Žliobaitė, I., Bifet, A., Pechenizkiy, M., & Bouchachia, A. (2014). A survey on concept drift adaptation. *ACM Computing Surveys*, 46(4), Article 44.
- Ghiselli, R. F., & Ismail, J. A. (1995). Gauging employee theft and other unacceptable behaviors in food service operations. *Hospitality Review*, 13(2), Article 3.
- Goldstein, M., & Uchida, S. (2016). A comparative evaluation of unsupervised anomaly detection algorithms for multivariate data. *PLOS ONE*, 11(4), e0152173.
- Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., & Pedreschi, D. (2018). A survey of methods for explaining black box models. *ACM Computing Surveys*, 51(5), Article 93.
- Haixiang, G., Yijing, L., Shang, J., Mingyun, G., Yuanyue, H., & Bing, G. (2017). Learning from class-imbalanced data: Review of methods and applications. *Expert Systems with Applications*, 73, 220 to 239.
- Ham, S., Kim, W. G., & Jeong, S. (2005). Effect of information technology on performance in upscale hotels. *International Journal of Hospitality Management*, 24(2), 281 to 294.

- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). Springer.
- He, H., & Garcia, E. A. (2009). Learning from imbalanced data. *IEEE Transactions on Knowledge and Data Engineering*, 21(9), 1263 to 1284.
- Hilal, W., Gadsden, S. A., & Yawney, J. (2022). Financial fraud: A review of anomaly detection techniques and recent advances. *Expert Systems with Applications*, 193, Article 116429.
- Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression* (3rd ed.). Wiley.
- Hospitality Financial and Technology Professionals. (2024). *Uniform system of accounts for the lodging industry* (12th rev. ed.). HFTP.
- Iglewicz, B., & Hoaglin, D. C. (1993). *How to detect and handle outliers*. ASQC Quality Press.
- Ilodigwe, L., & Adesemoye, A. C. (2021). The data backbone of health system transformation: A governance and architecture framework for interoperability, data quality, and advanced analytics at national scale. *International Journal of Health and Pharmaceutical Research*, 6(2), 52 to 76. <https://doi.org/10.56201/ijhpr.vol.6.no2.2021.pg52.76>
- 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 to 20.
- Ivanov, S., & Zhechev, V. (2012). Hotel revenue management: A critical literature review. *Tourism*, 60(2), 175 to 197.
- Julianty, F., Rapina, Setiawan, S., & Ndaru, H. P. S. (2022). Fraud in accounting: Surveys on hotels in Bandung. *Jurnal Riset Akuntansi Kontemporer*, 14(1), 75 to 82. <https://doi.org/10.23969/jrak.v14i1.4932>
- Kaufman, S., Rosset, S., Perlich, C., & Stitelman, O. (2012). Leakage in data mining: Formulation, detection, and avoidance. *ACM Transactions on Knowledge Discovery from Data*, 6(4), Article 15.
- Kimes, S. E. (1989). Yield management: A tool for capacity-constrained service firms. *Journal of Operations Management*, 8(4), 348 to 363.
- Kimes, S. E. (2008). The role of technology in restaurant revenue management. *Cornell Hospitality Quarterly*, 49(3), 297 to 309.
- King, G., & Zeng, L. (2001). Logistic regression in rare events data. *Political Analysis*, 9(2), 137 to 163.
- Kogan, A., Alles, M. G., Vasarhelyi, M. A., & Wu, J. (2014). Design and evaluation of a continuous data level auditing system. *Auditing: A Journal of Practice & Theory*, 33(4), 221 to 245.
- 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 to 122.
- Komi, N. M., & Adamolekun, A. (2021). Interpretable machine learning for early failure prediction in distributed renewable energy assets. *International Journal of Multidisciplinary Research and Growth Evaluation*, 2(6), 1015 to 1038. <https://doi.org/10.54660/IJMRGE.2021.2.6.1015-1038>
- Krawczyk, B. (2016). Learning from imbalanced data: Open challenges and future directions. *Progress in Artificial Intelligence*, 5(4), 221 to 232.
- Künsch, H. R. (1989). The jackknife and the bootstrap for general stationary observations. *The Annals of Statistics*, 17(3), 1217 to 1241.

- Ladapo, O. O., Dosunmu, A. A., Jooda, D., & Abolaji, T. O. (2022). Human-in-the-loop machine learning: A state of the art. *Journal of Frontiers in Multidisciplinary Research*, 3(1), 656 to 669. <https://doi.org/10.54660/JFMR.2022.3.1.656-669>
- Ladapo, O. O., Jooda, D., Dosunmu, A. A., & Abolaji, T. O. (2024). Keeping humans in the loop: Human-centered automated annotation with generative AI. *International Journal of Multidisciplinary Futuristic Development*, 5(1), 81 to 95. <https://doi.org/10.54660/IJMF.D.2024.5.1.81-95>
- Law, R., Leung, R., & Buhalis, D. (2009). Information technology applications in hospitality and tourism: A review of publications from 2005 to 2007. *Journal of Travel & Tourism Marketing*, 26(5 to 6), 599 to 623.
- Lawal, O. A., & Oduleye, T. E. (2019). Conceptualizing data driven executive decision systems for strategic financial planning. *Iconic Research and Engineering Journals*, 3(3).
- Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology*, 49(4), 764 to 766.
- Lipton, Z. C. (2018). The mythos of model interpretability. *ACM Queue*, 16(3), 31 to 57.
- Liu, F. T., Ting, K. M., & Zhou, Z.-H. (2008). Isolation forest. In *Proceedings of the 8th IEEE International Conference on Data Mining* (pp. 413 to 422).
- Liu, F. T., Ting, K. M., & Zhou, Z.-H. (2012). Isolation-based anomaly detection. *ACM Transactions on Knowledge Discovery from Data*, 6(1), Article 3.
- Lopez-Rojas, E. A., & Axelsson, S. (2014). BankSim: A bank payments simulator for fraud detection research. In *Proceedings of the 26th European Modeling and Simulation Symposium*.
- Lopez-Rojas, E. A., Elmir, A., & Axelsson, S. (2016). PaySim: A financial mobile money simulator for fraud detection. In *Proceedings of the 28th European Modeling and Simulation Symposium* (pp. 249 to 255).
- Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In *Advances in Neural Information Processing Systems 30* (pp. 4765 to 4774).
- Mbonu, I. S., Aliliele, C., Iwuanyanwu, U., & Uzoka, E. (2021a). A conceptual framework for risk based business intelligence architecture in financial technology platforms. *International Journal of Multidisciplinary Research and Growth Evaluation*, 2(6), 731 to 746. <https://doi.org/10.54660/IJMRGE.2021.2.6.731-746>
- Mbonu, I. S., Aliliele, C., Iwuanyanwu, U., & Uzoka, E. (2022a). A conceptual framework for AI enabled IT general controls and SOX audit automation processes. *Gyanshauryam, International Scientific Refereed Research Journal*, 5(5), 384 to 414. <https://doi.org/10.32628/GISRRJ2256239>
- Mbonu, I. S., Iwuanyanwu, U., Aliliele, C., & Uzoka, E. (2021b). A review of VoIP forensic analytics models for financial fraud detection and regulatory compliance monitoring. *International Journal of Multidisciplinary Research and Growth Evaluation*, 2(6), 711 to 730. <https://doi.org/10.54660/IJMRGE.2021.2.6.711-730>
- Mbonu, I. S., Iwuanyanwu, U., Aliliele, C., & Uzoka, E. (2022b). A review of data protection impact assessment models in multi cloud financial infrastructure systems. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 8(1), 589 to 623. <https://doi.org/10.32628/CSEIT25442>
- Mbonu, I. S., Iwuanyanwu, U., Uzoka, E., & Oluoha, O. M. (2019). Advances in enterprise log analytics and automated incident response architectures using Python and SIEM platforms.

- Iconic Research and Engineering Journals*, 3(2), 1000 to 1019.
<https://doi.org/10.64388/IREV3I2-1714915>
- Medon, J. J., & Oduleye, T. E. (2022). A comprehensive financial reporting model for strengthening compliance and organizational accountability systems. *International Journal of Multidisciplinary Research and Growth Evaluation*, 3(6), 768 to 777.
- Medon, J. J., & Oduleye, T. E. (2024). An integrated predictive analytics model for enhancing strategic financial forecasting and decision accuracy. *Gyanshauryam, International Scientific Refereed Research Journal*, 7(3), 258 to 279.
<https://doi.org/10.32628/GISRRJ247322>
- Mendes, T., Cardoso, P. J. S., Monteiro, J., & Raposo, J. (2022). Anomaly detection of consumption in hotel units: A case study comparing isolation forest and variational autoencoder algorithms. *Applied Sciences*, 13(1), Article 314.
<https://doi.org/10.3390/app13010314>
- Miller, T. (2019). Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence*, 267, 1 to 38.
- Moreno-Torres, J. G., Raeder, T., Alaiz-Rodríguez, R., Chawla, N. V., & Herrera, F. (2012). A unifying view on dataset shift in classification. *Pattern Recognition*, 45(1), 521 to 530.
- Morris, T. P., White, I. R., & Crowther, M. J. (2019). Using simulation studies to evaluate statistical methods. *Statistics in Medicine*, 38(11), 2074 to 2102.
- Nesvijejskaia, A., Ouillade, S., Guilmin, P., & Zucker, J. D. (2021). The accuracy versus interpretability trade-off in fraud detection model. *Data & Policy*.
<https://www.cambridge.org/core/journals/data-and-policy/article/accuracy-versus-interpretability-tradeoff-in-fraud-detection-model/3BA1586A6B635E08BE556FAB89AD8770>
- Ngai, E. W. T., Hu, Y., Wong, Y. H., Chen, Y., & Sun, X. (2011). The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature. *Decision Support Systems*, 50(3), 559 to 569.
- Niculescu-Mizil, A., & Caruana, R. (2005). Predicting good probabilities with supervised learning. In *Proceedings of the 22nd International Conference on Machine Learning* (pp. 625 to 632).
- Nigrini, M. J. (2012). *Benford's law: Applications for forensic accounting, auditing, and fraud detection*. Wiley.
- Nigrini, M. J., & Mittermaier, L. J. (1997). The use of Benford's law as an aid in analytical procedures. *Auditing: A Journal of Practice & Theory*, 16(2), 52 to 67.
- Nnabueze, S. B., Sakyi, J. K., Filani, O. M., Okojie, J. S., Abioye, R. F., Okereke, M., & Enow, O. F. (2024). Revenue optimization in energy distribution through integrated financial planning and advanced data-driven frameworks. *International Journal of Advanced Multidisciplinary Research and Studies*, 4(4), 1427 to 1445.
- Nnabueze, S. B., Sakyi, J. K., Filani, O. M., Okojie, J. S., Babatope, O. M., & Adanyin, A. (2023). Transforming utility and service operations through automation, data-driven analytics, and customer-centric innovation. *International Journal of Multidisciplinary Evolutionary Research*, 4(2), 40 to 57. <https://doi.org/10.54660/IJMER.2023.4.2.40-57>
- Nwakamma, S., Ojukwu, J., & Oyesiji, S. O. (2024). LLM-as-a-judge for automated model governance: A review of methods, biases, and release-gating practices. *World Journal of Innovation and Modern Technology*, 8(6), 185 to 216.
<https://doi.org/10.56201/wjimt.v8.no6.2024.pg185.216>

- Odeyemi, O. H., & Idowu, A. O. (2025). Financial sustainability of safety-net hospitals: A conceptual framework linking FinTech-enabled cost transparency, working-capital management, and access to capital. *International Journal of Technology, Health and Sustainability*, 1(2), 278 to 294.
- Oduleye, T. E., & Medon, J. J. (2021). A data-driven cost management model for improving strategic financial planning and performance evaluation. *International Journal of Multidisciplinary Research and Growth Evaluation*, 2(6), 524 to 537.
- Oduleye, T. E., & Medon, J. J. (2023). A predictive model for optimizing cash flow and working capital management in corporations. *Gyanshauryam, International Scientific Refereed Research Journal*, 6(5), 739 to 754. <https://doi.org/10.32628/GISRRJ236537>
- Ogbole, J. I., Okoruwa, P. O., Fadayomi, O., Akeju, B., Edivri, J., & Abolaji, T. O. (2025). Security analytics and digital forensics for enterprise risk management, advances and practical implications. *International Journal of Advanced Multidisciplinary Research and Studies*, 5(6), 2017 to 2028.
- Okafor, C. M., Farounbi, B. O., Adesanya, O. S., & Akinola, A. S. (2024). Controls for cross-border payments operations: Correspondent banking risk reduction via end-to-end monitoring. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 10(4), 1050 to 1071.
- Okoruwa, P. O., Babatope, O. M., Akokodaripon, D. A., & Akinleye, O. K. (2024). Developing integrated digital platforms for enhancing transparency in procurement and supply chain management. *International Journal of Multidisciplinary Research and Growth Evaluation*, 5(6), 1719 to 1729.
- Okoruwa, P. O., Fadayomi, O., Bello, A. D., Elebe, O., Hammed, N. I., & Omoegun, G. O. (2025). An artificial intelligence-driven financial crime investigation framework for analyst decision support. *International Journal of Multidisciplinary Research and Growth Evaluation*, 6(6).
- Olagunju, E. O., Edivri, J., & Oteri, O. (2023). A conceptual governance model for change, incident, and problem management in mission-critical enterprise IT environments. *Gyanshauryam, International Scientific Refereed Research Journal*, 6(6). <https://doi.org/10.32628/GISRRJ>
- Ominyi, M., & Anichukwueze, C. C. (2022). A conceptual framework for harmonizing data protection compliance across GDPR, NDPR, and United States privacy regimes. *International Journal of Social Sciences and Management Research*, 8(5), 148 to 174. <https://doi.org/10.56201/ijssmr.v8.no5.2022.pg148.174>
- Ominyi, M., & Anichukwueze, C. C. (2023). A review of algorithmic accountability and model risk management approaches in financial services. *Journal of Accounting and Financial Management*, 9(12), 218 to 238. <https://doi.org/10.56201/jafm.v9.no12.2023.pg218.238>
- Ominyi, M., & Anichukwueze, C. C. (2024). Conceptualizing responsible AI governance structures for automated decision systems in regulated industries. *International Journal of Social Sciences and Management Research*, 10(11), 403 to 430. <https://doi.org/10.56201/ijssmr.v10.no11.2024.pg.403.430>
- Ominyi, M., Anichukwueze, C. C., & Uzougbo, N. S. (2024). Developing a conceptual framework for AI-assisted compliance monitoring and anomaly detection in high-volume transaction environments. *World Journal of Innovation and Modern Technology*, 8(6), 204 to 228. <https://doi.org/10.56201/wjimt.v8.no6.2024.pg204.228>

- Oshoba, T. O., Hammed, N. I., & Odejebi, O. D. (2020). Blockchain-enabled compliance and audit trail model for cloud configuration management. *Journal of Frontiers in Multidisciplinary Research*, 1(1), 193 to 201.
- Oziri, S. T., Arowogbadamu, A. A. G., & Seyi-Lande, O. B. (2020). Predictive analytics applications in reducing customer churn and enhancing lifecycle value in telecommunications markets. *International Journal of Multidisciplinary Futuristic Development*, 1(2), 40 to 49.
- Ozowara, D. E., Adebayo, A., & Anunagba, C. O. (2025). Advanced conceptual model for strengthening audit quality using data analytics across financial institutions. *International Journal of Advanced Multidisciplinary Research and Studies*, 5(6), 2246 to 2268. <https://doi.org/10.62225/2583049X.2025.5.6.6050>
- Pang, G., Shen, C., Cao, L., & van den Hengel, A. (2021). Deep learning for anomaly detection: A review. *ACM Computing Surveys*, 54(2), Article 38.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825 to 2830.
- Perols, J. (2011). Financial statement fraud detection: An analysis of statistical and machine learning algorithms. *Auditing: A Journal of Practice & Theory*, 30(2), 19 to 50.
- Phua, C., Lee, V., Smith, K., & Gayler, R. (2010). *A comprehensive survey of data mining-based fraud detection research* (arXiv:1009.6119). arXiv.
- Pierce, L., Snow, D. C., & McAfee, A. (2015). Cleaning house: The impact of information technology monitoring on employee theft and productivity. *Management Science*, 61(10), 2299 to 2319.
- Politis, D. N., & Romano, J. P. (1994). The stationary bootstrap. *Journal of the American Statistical Association*, 89(428), 1303 to 1313.
- Provost, F., & Fawcett, T. (2001). Robust classification for imprecise environments. *Machine Learning*, 42(3), 203 to 231.
- Psychoula, I., Gutmann, A., Mainali, P., Lee, S. H., Dunphy, P., & Petitcolas, F. (2021). Explainable machine learning for fraud detection. *Computer*, 54(10), 49 to 59.
- Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?" Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 1135 to 1144).
- Rousseeuw, P. J., & Hubert, M. (2011). Robust statistics for outlier detection. *WIREs Data Mining and Knowledge Discovery*, 1(1), 73 to 79.
- Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, 1(5), 206 to 215.
- Saito, T., & Rehmsmeier, M. (2015). The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLOS ONE*, 10(3), e0118432.
- Sakyi, J. K., Filani, O. M., Nnabueze, S. B., Okojie, J. S., & Ogedengbe, A. O. (2022b). Developing KPI frameworks to enhance accountability and performance across large-scale commercial organizations. *Journal of Frontiers in Multidisciplinary Research*, 3(2), 81 to 93. <https://doi.org/10.54660/IJFMR.2022.3.2.81-93>
- Sakyi, J. K., Nnabueze, S. B., Filani, O. M., & Okojie, J. S. (2023). Revenue assurance strategies leveraging artificial intelligence and big data in service-intensive organizations.

- International Journal of Multidisciplinary Evolutionary Research*, 4(2), 58 to 75.
<https://doi.org/10.54660/IJMER.2023.4.2.58-75>
- Sakyi, J. K., Nnabueze, S. B., Filani, O. M., Okojie, J. S., & Okereke, M. (2022a). Customer service analytics as a strategic driver of revenue growth and sustainable business competitiveness. *Journal of Frontiers in Multidisciplinary Research*, 3(2), 109 to 123.
<https://doi.org/10.54660/IJFMR.2022.3.2.109-123>
- Sanni, J. O., & Wedraogo, L. (2024). Data-centric funnel optimization frameworks resolving revenue leakage in high-value services. *International Journal of Advanced Multidisciplinary Research and Studies*, 4(6), 2859 to 2874.
<https://doi.org/10.62225/2583049X.2024.4.6.5641>
- Sanni, J. O., Iwuanyanwu, U. A., & Essien, M. A. (2025). Problem-oriented process mining for auditable marketing automation lifecycle control. *International Journal of Advanced Multidisciplinary Research and Studies*, 5(6), 1933 to 1947.
<https://doi.org/10.62225/2583049X.2025.5.6.5650>
- Sanni, J. O., Iwuanyanwu, U. A., Essien, M. A., Atima, M. E., & Attah, A. (2022). Adaptive control models for AI-driven marketing automation in financial compliance environments. *Shodhshauryam, International Scientific Refereed Research Journal*, 5(1), 243 to 270.
<https://doi.org/10.32628/SHISRRJ247133>
- Sargent, R. G. (2013). Verification and validation of simulation models. *Journal of Simulation*, 7(1), 12 to 24.
- Schreyer, M., Sattarov, T., Borth, D., Dengel, A., & Reimer, B. (2017). *Detection of anomalies in large scale accounting data using deep autoencoder networks* (arXiv:1709.05254). arXiv.
- Seyi-Lande, O. B., Arowogbadamu, A. A. G., & Oziri, S. T. (2022). Cross-functional key performance indicator frameworks for driving organizational alignment and sustainable business growth. *International Journal of Multidisciplinary Futuristic Development*, 1(2), 1 to 18.
- Siddiqi, N. (2006). *Credit risk scorecards: Developing and implementing intelligent credit scoring*. Wiley.
- Talluri, K. T., & van Ryzin, G. J. (2004). *The theory and practice of revenue management*. Kluwer Academic.
- Thiprungsri, S., & Vasarhelyi, M. A. (2011). Cluster analysis for anomaly detection in accounting data: An audit approach. *The International Journal of Digital Accounting Research*, 11, 69 to 84.
- Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B*, 58(1), 267 to 288.
- Tonoyan, A., Dada, O., & Ayivi-Donkor, S. S. (2021a). Advances in demand forecasting: Machine learning algorithms for revenue projection and financial planning accuracy. *Gyanshauryam, International Scientific Refereed Research Journal*, 4(1), 282 to 317.
- Tonoyan, A., Dada, O., & Ayivi-Donkor, S. S. (2021b). Predictive cash flow modeling in retail: A review of advanced forecasting and working capital optimization. *Shodhshauryam, International Scientific Refereed Research Journal*, 4(3), 366 to 397.
- Tonoyan, A., Dada, O., & Ayivi-Donkor, S. S. (2022). Algorithmic process optimization and cost reduction: A review of simulation modeling and financial impact assessment. *Journal of Accounting and Financial Management*, 8(8), 139 to 169.
<https://doi.org/10.56201/jafm.vol.8.no8.2022.p139.169>

- Tonoyan, A., Dada, O., & Ayivi-Donkor, S. S. (2024). Real-time KPI tracking systems: A review of automated performance monitoring and data-driven decision making. *World Journal of Innovation and Modern Technology*, 8(6), 247 to 281. <https://doi.org/10.56201/wjimt.v8.no6.2024.pg247.281>
- Uchechi, N., Ominyi, M., Anichukwueze, C. C., & Jones, B. C. (2022). Real-time compliance pipelines: A review of automated risk monitoring and enterprise data governance. *International Journal of Multidisciplinary Futuristic Development*, 3(2), 66 to 94. <https://doi.org/10.54660/IJMFD.2022.3.2.66-94>
- Uchechi, N., Ominyi, M., Anichukwueze, C. C., & Jones, B. C. (2023). Advances in standardizing un-standardized compliance controls for high-volume non-traditional distribution networks. *International Journal of Advanced Multidisciplinary Research and Studies*, 3(6). <https://doi.org/10.62225/2583049X.2023.3.6.6838>
- Ustun, B., & Rudin, C. (2016). Supersparse linear integer models for optimized medical scoring systems. *Machine Learning*, 102(3), 349 to 391.
- Van Calster, B., McLernon, D. J., van Smeden, M., Wynants, L., & Steyerberg, E. W. (2019). Calibration: The Achilles heel of predictive analytics. *BMC Medicine*, 17, Article 230.
- Vasarhelyi, M. A., Kogan, A., & Tuttle, B. M. (2015). Big data in accounting: An overview. *Accounting Horizons*, 29(2), 381 to 396.
- Walawalkar, G., Adesuyi, M. O., Kalu, A., & Oduleye, T. E. (2025). Executive financial dashboards for real-time strategic oversight. *International Journal of Advanced Multidisciplinary Research and Studies*, 5(6), 2042 to 2054.
- Weatherford, L. R., & Kimes, S. E. (2003). A comparison of forecasting methods for hotel revenue management. *International Journal of Forecasting*, 19(3), 401 to 415.
- West, J., & Bhattacharya, M. (2016). Intelligent financial fraud detection: A comprehensive review. *Computers & Security*, 57, 47 to 66.
- Wibawa, I. M. S. T., Karyawati, A. A. I. N. E., Putri, L. A. A. R., & Santiyasa, I. W. (2025). Development of an information system with a machine learning approach for anomaly detection in hotel transactions. *Jurnal Elektronik Ilmu Komputer Udayana*, 14(1).