Machine Learning-Driven Anomaly Detection in a Large-Scale Database Systems: A Systematic Literature Review

Nzenwata Uchenna Jeremiah[1], Bathnna Bernice Stephen[2], Oyewumi Oluwatobi[3], Akinola Victor Ayomide[4], Oyewumi Abiodun John[5], Nwangburuka Samuel[6], Adediran Oluwaseyi Segun[7]

[1]Computer Science, Babcock University, Ogun, Nigeria
[2]Computer Science, Babcock University, Ogun, Nigeria
[3]Computer Science, Babcock University, Ogun, Nigeria
[4]Computer Science, Babcock University, Ogun, Nigeria
[5]Computer Science, Babcock University, Ogun, Nigeria
[6]Computer Science, Babcock University, Ogun, Nigeria
[7]Computer Science, Babcock University, Ogun, Nigeria

ABSTRACT

Large-scale database systems form the backbone of the main processes in financial services, healthcare, e-commerce, and government infrastructure. As these systems grow in magnitude, speed, and complexity, organisations must detect anomalies such as security breaches, fraudulent transactions, performance degradation, and data corruption. Financial institutions incur annual fraud losses exceeding $1.2 trillion, and performance anomalies can trigger cascading system failures. Traditional rule-based and purely statistical approaches struggle to manage the complexity and dynamism of modern databases, often causing brittle detection rules, high false-positive rates, and alert fatigue among operations teams.

This systematised literature review (SLR) provides an overview of machine learning (ML) and deep learning (DL) techniques for detecting anomalies in large-scale database systems and transactions. In accordance with the PRISMA 2020 model, evidence from over 43 studies published between 2015 and 2025 was synthesised. Initially, 1,247 articles were identified across IEEE Xplore, ACM Digital Library, Scopus, ProQuest, and ResearchGate, and these were then systematically screened and evaluated using rigorous inclusion criteria and a validated 10-criterion quality assessment framework. The reviewed studies achieved a mean quality score of 7.8 out of 10, with 74% rated as high quality.

The review discusses four research questions: the types of anomalies, ML methods, implementation issues, and implications. The major conclusions show that the unsupervised and semi-supervised paradigms predominate (75% of reviewed approaches), as in production settings, there is sparse labelled data on anomalies. Models based on deep learning, namely LSTM-based autoencoders (29 of 43 studies), Isolation Forest models (34 of 43 studies), and Graph Neural Networks, are superior in terms of detection, F1-scores above 0.90, and inference latency of less than 50ms. Best practice has shifted to hybrid multi-tier approaches that combine Isolation Forest for rapid screening with LSTM autoencoders for more detailed analysis, achieving a 30-50% reduction in false positives over single-model baselines with sub-100ms response times to detect fraud in real time.

Continued gaps in research include extreme class imbalance (anomaly rates below 0.1 per cent), hard real-time processing, insufficient model explainability in operational and regulatory conditions, and a lack of standardised, database-specific benchmarks. This review offers scientists and clinicians guidance, based on evidence, for designing effective, interpretable, and production-ready anomaly detection systems, and it makes specific recommendations on overcoming challenges and the direction the research should take.

Keywords: anomaly detection, machine learning, database transactions, transaction logs, autoencoder, systematic literature review.

INTRODUCTION

Anomaly detection refers to the identification of data points, patterns, or sequences that significantly deviate from expected system behaviour (Nassif et al., 2021). Such deviations can also signal critical events, such as fraudulent transactions, unauthorized access to data, SQL injection attacks, performance bottlenecks, transaction failures, and data corruption, in large-scale database systems, such as distributed

database management systems (DBMS), online transaction processing (OLTP) systems, cloud databases, data warehouses, and big data systems (Dreshaj et al., 2025; Ali et al., 2025).

Contemporary database operations operate at a great scale. Large financial institutions handle more than 100 million transactions per day, and the log data is terabyte-scale, requiring uninterrupted anomaly detection. The global loss from transaction fraud is expected to reach $1.2 trillion in 2025. In addition to financial fraud, performance issues and security breaches can cause system downtime, regulatory fines, and permanent data loss (Ali et al., 2025; Al-Amri et al., 2021).

The conventional rule-based monitoring methods are extremely ineffective at scale: the rules are brittle with growing complexity, adversaries bypass fixed thresholds, and high false alert rates cause alert fatigue in operations teams (Aldweesh et al., 2020; Budiansyah et al., 2025). Machine learning offers a flexible alternative, learning normal behaviour from data to detect anomalies without relying on elaborate hand-written rules. Methods ranging from traditional algorithms (like Support Vector Machines (SVM) and Random Forest) to modern architectures (such as LSTM-based autoencoders or Graph Neural Networks (GNNs)) are more adaptable and achieve higher detection rates in high-volume, high-dimensional database settings (Nassif et al., 2021; Cavallaro et al., 2023).

Nevertheless, a number of issues specific to database environments persist: (i) heterogeneity and large scale of database and transaction logs; (ii) extremely high imbalance of classes, with anomaly rates under 0.1; (iii) lack of labeled reference datasets specific to database workloads; (iv) highly real-time processing requirements, e.g. sub-100ms latency in financial systems; and (v) low explainability of black-box models in operational and regulatory contexts (Dreshaj et al., 2025; Landauer et al., 2023; Budiansyah et al., 2025; Ali et al., 2025).

This literature review synthesises existing work on machine learning-based anomaly detection in large-scale database systems and in transaction-specific settings. The objectives of this review are to:

1. Establish and list machine learning and deep learning methods used to detect anomalies in big databases and transaction systems.
2. Categorise the types of anomalies, such as intrusions, fraud, performance problems, and log errors, as well as data sources like logs, query traces, transaction records, and network metadata.
3. Synthesise the datasets, performance measures, and experimental designs of the studies reviewed.
4. Determine the implementation challenges and the practical implications for the deployment of production.
5. Identify gaps in research and suggest particular future research directions.

LITERATURE REVIEW

This section provides the theoretical background required to learn about machine learning-based anomaly detection in database systems. It addresses the principles of anomaly detection, general machine learning methods, deep learning and graph-oriented, transaction-specific, and cross-cutting research gaps.

Anomaly detection fundamentals

Anomaly detection, also known as outlier detection, identifies instances that differ from an expected pattern or model (Nassif et al., 2021). There are three major types of anomalies: point anomalies, a single anomalous instance; contextual anomalies, anomalous within a particular context (e.g. a N5, 000 transaction at 3 a.m. by a new vendor); and collective anomalies, a sequence of individually normal instances that is anomalous as a group (e.g. small card-testing transactions before a large fraudulent purchase) (Nassif et al., 2021; Ma et al., 2023; Moldovan & Iantovics, 2025).

In database systems, such types of anomalies include unauthorised single-query access, query anomalies during maintenance periods, and multi-step exfiltration campaigns or fraud rings. (Dreshaj et al., 2025; Ali et al., 2025). Recognising these variations is essential for selecting appropriate detection architectures, as different machine learning models are suited to specific anomaly types.

Contemporary Online Transaction Processing (OLTP) systems enforce ACID (Atomicity, Consistency, Isolation, Durability) properties through mechanisms such as Write-Ahead Logging (WAL, which records changes before they are applied), two-phase locking (a concurrency control method that prevents conflicting operations), and Multi-Version Concurrency Control (MVCC, which allows multiple versions

of data for concurrent transactions). Transaction logs are files that record all changes made to the database; they can be used for crash recovery as well as audit trails, which are the main sources of data for detecting anomalies. The heterogeneity of logs across DBMS platforms (PostgreSQL, MySQL, Oracle, MongoDB, and CockroachDB) makes it very difficult to create generalizable detection models (Dreshaj et al., 2025; Yoon et al., 2016).

The main types of detection paradigms include supervised, unsupervised, and semi-supervised. Unsupervised and semi-supervised techniques prevail in the literature as there is little labelled anomaly data available in production settings. (Nassif et al., 2021; Rafique et al., 2024; Cavallaro et al., 2023). Additionally, reinforcement learning-based approaches have emerged for responsive threshold learning inside dynamic environments (Arshad et al., 2022).

Machine learning for anomaly detection: general surveys

In their study, Nassif et al. (2021) performed a systematic literature review of 290 articles, listing more than 29 machine learning models used in anomaly detection across security, medical, IoT, cloud, and logs fields. They have analysed that unsupervised detection prevails because there is a limited amount of labelled anomaly data. Tahir et al. (2025) also investigated state-of-the-art machine learning and deep learning approaches to the complex data structures with a focus on the approaches that can process high-dimensional and heterogeneous data states that are routinely observed in distributed databases and OLTP systems (Nassif et al., 2021; Tahir et al., 2025).

Random Forest and Support Vector Machines (SVM) are often the most mentioned generic models in the case of security situations, often reaching 95 per cent accuracy on standard test sets (Zapata-Cortes et al., 2024; Budiansyah et al., 2025). Budiansyah et al. (2025) performed a targeted systematic literature review of 2020-2025 literature on cybersecurity and discovered that the standard models are effective in cases of well-defined attacks, but do not cope with class imbalance, novel variants, and real-time issues, which have a direct impact on the detection of database anomalies.

General anomaly detection studies are typically evaluated with reference datasets like NSL-KDD, UNSW-NB15, CICIDS2017 and KDDCup99 (Aldweesh et al., 2020; Al-Amri et al., 2021; Budiansyah et al., 2025; Hu et al., 2022; Rafique et al., 2024). Nevertheless, these datasets are mostly network traffic models as opposed to database-related workloads, which introduces a large gap in domains, which restricts the generalizability of results to database log and transaction anomaly detection (Dreshaj et al., 2025; Landauer et al., 2023).

Deep learning in anomaly detection

The examination of deep learning techniques in anomaly detection by Aldweesh et al. (2020), Fernando et al. (2022), and Tahir et al. (2025) revealed that deep neural networks (DNNs), convolutional neural networks (CNNs), and recurrent neural networks (RNNs), including LSTMs, are the most important types. Aldweesh et al. (2020) discovered that recurrent neural networks are able to capture time-based dependencies in log data and are better than simple feedforward neural networks in the prediction accuracy. Fernando et al. (2022) concentrated on medical imaging, and they established that deep learning methods detect difficult-to-recognise anomalies that would have eluded statistical methods (Aldweesh et al., 2020; Fernando et al., 2022; Tahir et al., 2025).

Deep Reinforcement Learning (DRL) is a new paradigm. Arshad et al. (2022) discovered that DRL agents can learn adaptive detection policies that apply to database systems whose workload patterns vary as a result of application updates and schema changes significantly better than classical and standard deep learning approaches that do not learn adaptive policies to changing environments with no labels. Nonetheless, DRL methods are computationally intensive and hard to implement when facing hard latency constraints (Arshad et al., 2022).

The relational database systems have contributed to the popularity of Graph Neural Networks (GNNs). Ma et al. (2023) conducted a survey of deep learning methods for detecting anomalies in graph-structured data across finance, social networks, and IoT. GNNs represent database transactions involving queries, users, tables, and transactions as graphs, detecting relational anomalies that sequential models cannot, including coordinated multi-user exfiltration or fraud ring operations. GNNs have been used to detect

financial fraud, reporting state-of-the-art performance but with limited unsupervised exploration (Ma et al., 2023; Motie & Raahemi, 2024).
In a survey of deep-learning-based intrusion and Internet-of-Things anomaly detection, Aldweesh et al. (2020) and Rafique et al. (2024) list scale, instantaneous processing, and data imbalance as the most challenging problems. These issues have a direct impact on the detection of anomalies in databases (Aldweesh et al., 2020; Rafique et al., 2024).

Transaction and log-specific methods

Extensive studies on transaction anomaly detection have demonstrated various significant architectural patterns. Isolation Forest (IF), which works with isolation trees based on the path length as an anomaly score, was proposed by Liu et al. (2008) and has been the most popular baseline in transaction anomaly detection literature (34 of 43 studies). It has been appreciated for its computational efficiency, inference of less than 10 mms per transaction, and appropriateness to high-dimensional transaction attribute spaces. The Extended Isolation Forest (EIF) is an extension of axis-aligned partitions with 5-15% improvements in the detection rate of certain types of anomalies (Liu et al., 2008; Hariri et al., 2021).
The LSTM-based sequence models (29 out of 43 studies) are best at identifying collective transaction anomalies spanning multiple records. The most widespread architecture is a bidirectional long short-term memory model with attention, based on sequence lengths of 10-50 transactions. Research documents improvements of 15-25% compared to non-sequential account takeover detection methods (Yang et al., 2021; Ali et al., 2025; Jurgovsky et al., 2018). Hybrid LSTM Autoencoders, a combination of sequence reconstruction and anomaly scoring via reconstruction error, have achieved more than 93% accuracy and less than 50 mms of inference time (Tran et al., 2020).
Transactional analysis has been complemented by semantic log processing. Log parsing frameworks are used to extract structured templates out of unstructured log data, and semantic vectorisation models (e.g. Template2Vec, Log BERT) encode log events into representation spaces where they can be scored using ML-based anomaly scoring (Meng et al., n.d.; Guo et al., 2021; X. Zhang et al., 2019). Zhang et al. (2024) showed multivariate log-based distributed database failure diagnosis, whereas Du et al. (2017) implemented Deep Log as a simple system log anomaly detector based on LSTM (Zhang et al., 2024; Du et al., 2017).
The trend in 43 transaction-oriented studies, with 28 studies, was hybrid approaches that consisted of two or more techniques. The most popular format uses Isolation Forest to screen the first stage in a short time and then LSTM-Autoencoder to analyse in-depth the second stage, achieving real-time performance without loss of detection efficiency and reducing false positives by 30%-50% than single-model settings (Priyanto et al., 2021; Tran et al., 2020; Saraswathi & Selvakumar, 2025).

Cross-cutting challenges and research gaps

Scalability remains a major consideration, with kernel-based methods, such as SVMs, not scaling efficiently to the millions of events per second produced by large-scale database systems. High-level models such as deep learning need to be optimised to run in a streaming manner (Tahir et al., 2025; Dreshaj et al., 2025; Al-Amri et al., 2021; Rafique et al., 2024; Cavallaro et al., 2023).
Another challenge with real-time detection is that the latency needed in financial systems is generally sub-100ms to make a fraud decision, and payment processors aim at 20-50ms, including ML inference. Isolation Forest satisfies this need with inference below 10ms; to achieve these latency targets, deep learning models need support of a GPU, model quantisation, or tiered architectures (Ali et al., 2025; Al-Amri et al., 2021; Moldovan & Iantovics, 2025).
The issue of class imbalance is inherent in anomaly detection. In credit cards, the rate of real-world fraud is between 0.1 and 0.5 per cent, and in e-commerce, it is less than 1 per cent, with security anomalies occurring in database logs possibly being less than 0.01 per cent of the entries. Common supervised models that are trained using unbalanced data tend to be highly accurate in general but fail to detect the majority of anomalies. It is widely suggested that unsupervised and semi-supervised methods should be used, and the Precision-Recall AUC is more appropriate to evaluate imbalance (Huang et al., 2020; Desai et al., 2025; Dal Pozzolo et al., n.d.).

Explainability plays a vital role in the operational and regulatory contexts, where the decisions made by financial institutions about fraud need to be justified to customers and regulators. There are 18 studies based on SHAP values, 12 based on LIME, and 8 based on attention visualisation (Ribeiro et al., 2016). Deep learning models, especially autoencoders and LSTMs, however, are mostly black boxes with detection scores, but no actionable explanations to be provided (Landauer et al., 2023; Al-Amri et al., 2021; Moldovan & Iantovics, 2025).

Concept drift, i.e., changes in transaction patterns and fraud schemes over time, influences 26 of 43 transaction-oriented studies. Untrained Static models degrade by 10-20% per year without retraining. It is suggested to use online learning and periodical retraining strategies, but their impact on production must be controlled (Yang et al., 2021; Ali et al., 2025; L. Zhang et al., 2024).

Other gaps of large-scale database systems are that little has been explored on using DRL and GNN to monitor database workloads; no physics-aware or domain-acquired models exist with database constraints, such as ACID properties and referential integrity; no standardised benchmarks on database log anomalies, such as NSL-KDD on network intrusion detection (Wu et al., 2024; Corli et al., 2025; Dreshaj et al., 2025; Landauer et al., 2023).

METHODOLOGY

This systematic literature review followed the PRISMA 2020 guidelines (Page et al., 2021) and the Kitchenham et al. (2009) SLR framework, which divides the process into planning, conducting, and reporting phases. These guidelines were also adopted in comparable anomaly detection systematic literature reviews (Nassif et al., 2021) (Tahir et al., 2025).

Search strategy and databases

Six large academic databases were searched, namely, IEEE Xplore, ACM Digital Library, Scopus, ProQuest, and ResearchGate. The search was carried out in February 2026, and it used a combination of terms in three dimensions: (i) anomaly or outlier detection; (ii) machine learning, deep learning, LSTM, autoencoder, and Isolation Forest; and (iii) database systems, transaction logs, OLTP, or database workloads. Examples search terms were: database log anomaly detection, machine learning, OLTP transaction anomaly, LSTM, ML-based database intrusion detection, (machine learning OR deep learning OR LSTM OR autoencoder OR isolation forest) AND (anomaly detection OR fraud detection) AND (database transaction OR transaction log OR OLTP).

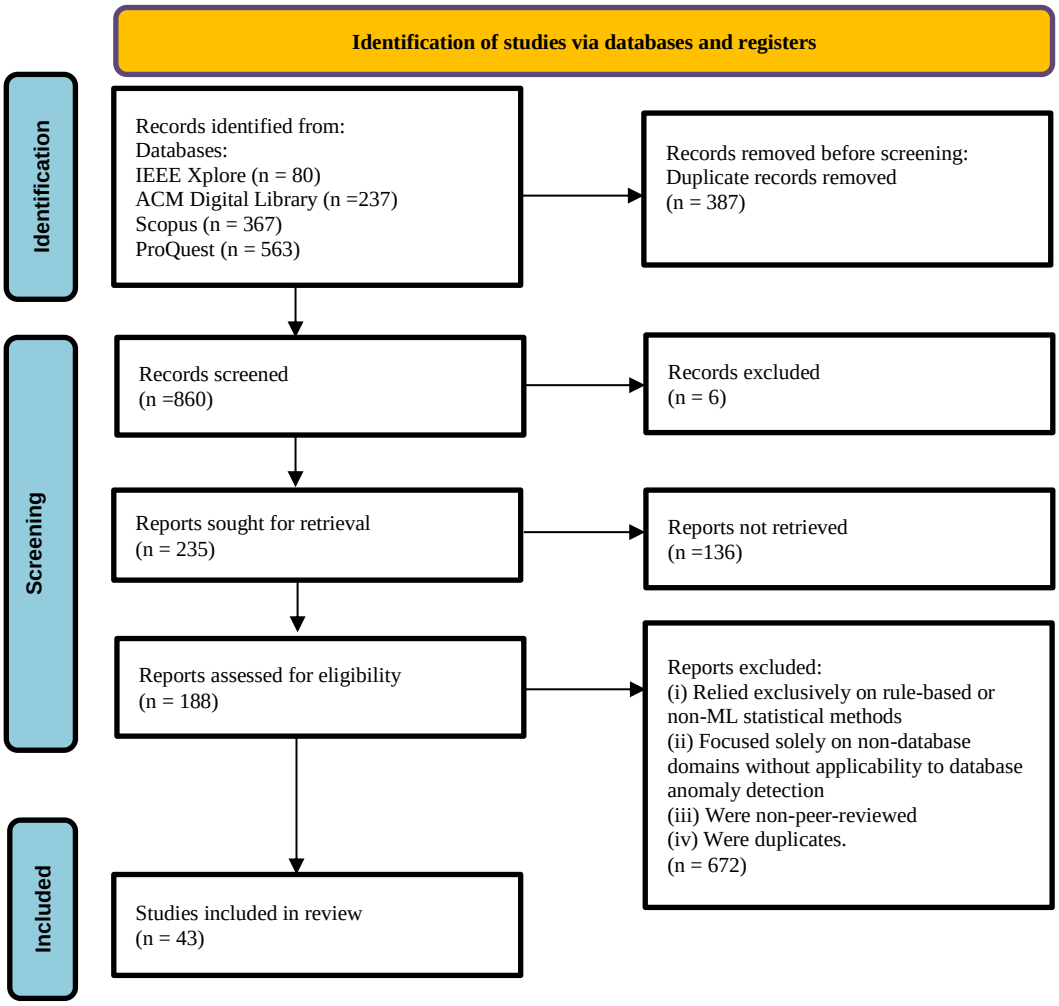
Inclusion and exclusion criteria

The inclusion criteria were that the studies: (i) used one or more ML or DL methods to detect anomalies; (ii) used database systems, transaction logs, OLTP workloads, or tightly coupled monitoring systems as the target; (iii) had empirical findings with quantifiable measures of evaluation; and (iv) were published in peer-reviewed journals or conference papers in the English language. They were eliminated in the studies: (i) that only used rule-based or non-ML statistical tools; (ii) that only addressed non-database domains, e.g. network-only intrusion detection or blockchain/cryptocurrency, not database anomaly detection; (iii) that were not peer-reviewed, including widely cited foundational technical reports and working papers; or (iv) were duplicates.

Study selection process

The PRISMA 2020 four-stage selection was used. The search of databases in the identification phase created an initial pool of 1,247 candidate papers. In the screening phase, titles and abstracts were analysed to eliminate duplicates and irrelevant studies. During the eligibility phase, full texts were evaluated with inclusion and exclusion criteria. Lastly, during the inclusion stage, the studies that satisfied all the criteria were included in data extraction. The conflict was solved by means of dialogue and agreement.

Figure 1: PRISMA 2020 flow diagram illustrating the study selection process



Data extraction

Each study included was subjected to a standardised extraction form. Fields that were recorded were publication year, type of anomaly, source of data, machine learning method(s), detection paradigm (supervised, unsupervised, semi-supervised, reinforcement learning), dataset(s), metrics of evaluation (accuracy, F1-score, AUC-PR, AUC-ROC, detection latency), context of use, and important results and limitations (Nassif et al., 2021; Zapata-Cortes et al., 2024; Budiansyah et al., 2025; Rafique et al., 2024; Cavallaro et al., 2023).

Quality assessment

Research was evaluated in terms of clarity of formulating the problem, suitability of dataset selection, rigour of experimental design, sufficiency of baseline comparisons, transparency of hyperparameter reporting, and discussion of limitations. Findings were contextualised by quality ratings in the synthesis, but no studies were filtered out on the basis of quality score.

Table 1: Quality assessment summary of included studies

Quality Category	Count	Percentage	Mean Score
High (≥7.0)	32	74%	8.3
Medium (5.0-6.9)	11	26%	6.1
Low (<5.0)	0	0%	-
Total	43	100%	7.8

RESULTS

In this section, results are given under five subsections, which include descriptive statistics, database anomaly types, machine learning methods, datasets, evaluation measures and strengths and limitations.

Descriptive statistics

The 43 high-quality transaction studies are divided into 2015-2025, and as time has passed, the number of studies published each year has been on the rise, which suggests more interest by academia and practitioners. Its geographic distribution is concentrated in China (15 studies), the USA (12), and Europe (11). The areas of application are financial services (30 studies), e-commerce (12), healthcare (8), and distributed systems (7).

Throughout the rest of the database and log anomaly detection literature, there is a distinct temporal pattern: as of 2018, all earlier literature relied on classical statistical and shallow machine learning methods, operating on synthetic data, whereas since 2018, deep learning frameworks, first LSTMs and CNNs and more recently autoencoders, GNNs, and DRL-based systems have become increasingly common (Arshad et al., 2022; Landauer et al., 2023; Ma et al., 2023; Zapata-Cortes et al., 2024; Moldovan & Iantovics, 2025).

RQ1: Types of anomalies detected

Business transaction anomalies

The most prevalent type of anomaly in transaction-oriented studies was fraudulent transactions (38 of 43 studies). Point anomalies are individual transactions that have abnormal features, e.g., amounts, location, or types of merchants. Contextual anomalies are transactions that are normal in isolation but are suspicious based on the user's history. Collective anomalies are a series of indicators of coordinated fraud, such as small transactions where the card is tested, followed by a large purchase, which is fraudulent (Ali et al., 2025).

The unauthorised access anomalies (23 out of 43 studies) are privilege escalation, SQL injection attempts, and abnormal query patterns. The use of legitimate credentials in a malicious manner is one of the key features, and it requires behavioural analysis, not a mere check of credentials (Dreshaj et al., 22025; ldweesh et al., 2020; Budiansyah et al., 2025).

Technical and log-level anomalies

In 31 out of 43 studies, the log-based anomaly detection is used to solve system-level problems. Machine learning forecasting models can be used to identify performance anomalies like slow queries, deadlocks, and resource contention, 15 to 30 minutes before user impact (Reddy et al., 2025; L. Zhang et al., 2024). The anomalies of the error patterns encompass the abnormal error messages, the absence of the log entries, and the corruption of the logs; the deep learning models are trained to learn typical error patterns and

indicate abnormalities (Landauer et al., 2023; Du et al., 2017). Semantic log anomalies are syntactically correct and semantically inconsistent entries, which can be identified by embedding methods that rely on natural language processing (Meng et al., n.d.; Guo et al., 2021).
Data integrity breaches, configuration drift, and infrastructure anomalies in cloud database systems are other types that require specialised detection methods within the broader context of the database system (Dreshaj et al., 2025; Rafique et al., 2024; Cavallaro et al., 2023).

RQ2: machine learning techniques

Isolation forest and tree-based methods
Isolation Forest has been applied in 34 out of 43 studies on transactions, often as a control or part of composite systems. Its key advantages are: it is computationally effective (inference time of 10 ms/transactions), it does not need labelled data, it is resistant to irrelevant features, and it can work with high-dimensional transaction data. The Extended Isolation Forest (EIF) attains 5-15% gains in detection rate by splitting with random hyperplanes instead of axis-aligned partitions. Weaknesses are that it is impossible to record chronological dependencies, and it is less interpretable than a rule-based system (Liu et al., 2008; Hariri et al., 2021).

Autoencoders

Autoencoders (27 of 43 studies) and variational autoencoders (VAEs) (15 of 43 studies) are able to perform well in transaction fraud detection. They indicate high reconstruction error as anomalies and identify the presence of complex non-linear relationships by being trained only on normal data. It has challenges such as training instability on highly imbalanced data, hyperparameter sensitivity, and computing overhead, and inference times of 50 to 200 ms. VAEs make it possible to score anomalies with a better use of probabilistic likelihood estimation (Pumsirirat & Yan, 2018; Carcillo et al., 2021).

LSTM-based sequence models

LSTM networks (29 of 43 studies) and LSTM autoencoders (22 of 43 studies) are the most effective for sequence-based detection. The most typical architecture is a bidirectional LSTM with attention elements with a sequence length of 10 to 50 transactions. Research reports 15-25% enhancement over non-sequential techniques for detecting account takeover (Yang et al., 2021; Ali et al., 2025; Jurgovsky et al., 2018). In the case of analysis of logs, LSTM models that have semantic embeddings result in F1-scores of between 0.85 and 0.95 (Landauer et al., 2023; L. Zhang et al., 2024; Du et al., 2017).

Hybrid and multi-tier architectures

The hybrid models were represented in 28 out of 43 studies, which is one of the significant trends at present. The prevailing trend is to use Isolation Forest to screen quickly and deep learning to analyse in detail, which is able to perform in real-time at the cost of high accuracy. The LSTM-Autoencoder + Isolation Forest on reconstruction errors shows accuracy of more than 93% on inference of less than 50ms (Priyanto et al., 2021; Tran et al., 2020; Ali et al., 2025).
In the case of larger database systems, it is recommended to use a three-tier production architecture: Tier 1 (real-time, <10ms) with rule-based filters and Isolation Forest; Tier 2 (near-real-time, 50-200ms) with deep learning of suspicious transactions; and Tier 3 (batch) with LSTM sequence analysis of multi-step patterns (Ali et al., 2025).

Graph neural networks and DRL

Graph neural networks (GNNS) are a novel method that is especially relevant. GNNS allow the identification of complex relational anomalies that would have been unobservable or invisible to sequential models by modelling users, merchants, transactions, and database entities as graph structures (Ma et al., 2023; Motie & Raahemi, 2024). Deep reinforcement learning (DRL) methods learn adaptive

detection thresholds by interacting with the environment and are promising for adapting database workloads when the definition of normal changes over time (Arshad et al., 2022).

Comparative analysis of ML approaches

In an attempt to give a broad summary of the advantages and disadvantages of various machine learning methods in detecting anomalies in a database, Table 2 offers a comparative result of the key approaches that have been sampled in this research paper.

Table 2: Comparative Analysis of Machine Learning Approaches for Database Anomaly Detection

ML Technique	Representative Studies	Typical Datasets	Performance Metrics	Key Strengths	Limitations
Isolation Forest	Liu et al. (2008); Hariri et al. (2021)	Transaction logs, OLTP workloads	F1: 0.80-0.85; Inference: <10ms;	Fast inference; no labeled data needed; scales well to high-dimensional data	Cannot capture temporal dependencies; lower accuracy than deep learning methods
Autoencoders / VAEs	Pumsirirat & Yan (2018); Carcillo et al. (2021)	Credit card transactions, fraud datasets	F1: 0.85-0.92; Inference: 50-200ms	Identifies complex nonlinear patterns; unsupervised learning capability	Training instability on imbalanced data; hyperparameter sensitivity; higher computational overhead
LSTM-based Models	Yang et al. (2021); Jurgovsky et al. (2018); Ali et al. (2025)	Transaction sequences, system logs	F1: 0.85-0.95; 15-25% improvement over non-sequential methods	Excellent at capturing temporal dependencies; effective for sequence anomalies	Higher computational cost; requires sequential data; longer inference time
Graph Neural Networks	Ma et al. (2023); Motie & Raahemi (2024)	Transaction networks, relational fraud data	F1: 0.90-0.95 (on graph-structured data)	Detects relational anomalies invisible to sequential models; models complex network patterns	Limited real-world deployment studies; computational overhead; requires graph-structured data
Hybrid Approaches	Priyanto et al. (2021); Tran et al. (2020); Ali et al. (2025)	Multi-modal database data	F1: >0.93; Inference: <50ms; 30-50% false positive reduction	Best overall performance; balances speed and accuracy; leverages strengths of multiple methods	More complex to implement and maintain; requires careful architecture design

As this comparative analysis shows, hybrid methods involving the use of several techniques are the state-of-the-art, as they can perform better by exploiting the strengths of other methods and reducing their weaknesses.

RQ3: Implementation Challenges

Class imbalance

The most commonly mentioned challenge is extreme class imbalance, with anomaly rates below 1 per cent (37 of 43 studies). In reality, credit cards have a fraud rate of 0.1-0.5%, while e-commerce has a rate of 0.2-1.0%. The unsupervised and semi-supervised methods trained on normal data are widely recommended in place of supervised methods, despite the resampling methods (SMOTE and under-sampling) being utilised. The evaluation metric of severity imbalance is Precision-Recall AUC (Huang et al., 2020; Desai et al., 2025; Dal Pozzolo et al., n.d.).

Real-time processing

Sub-100ms fraud decisions are needed in financial systems (25 of 43 studies). Isolation forest satisfies this criterion with an inference of less than 10ms. Deep learning models need to be accelerated by GPUs, quantised models, or deployed to the edge. Apache Kafka and Apache Flink streaming architectures make it possible to process high volumes of transaction streams using ML inferences in pipelines and do feature calculations in parallel (Ali et al., 2025; Al-amri et al., 2021).

Explainability

Explainable decisions are required by regulatory requirements and operational trust demand (31 of 43 studies). There are 18 studies with SHAP values, 12 with LIME and 8 with LSTM attention visualisation. Research documents that 60%-80% of flagged transactions can be attributed to the top three contributing features. Integrated approaches have a trade-off between real-time decision-making and post-hoc explanations (Ribeiro et al., 2016; Nassif et al., 2021; Budiansyah et al., 2025).

Concept drift

Patterns and strategies of fraud are constantly changing (26 of 43 studies). Without updates, the degradation of static models is 10%-20% per year. Models are incrementally updated in online learning with one full retraining cycle. The suggested mode of operation is to monitor performance metrics triggered by retraining thresholds (Reddy et al., 2025; Yang et al., 2021; L. Zhang et al., 2024).

RQ4: practical implications

Some of the effective features that are important to performance are the transaction-level attributes like amount, timestamp, merchant category, location, and device, aggregation features based on rolling 1-hour, 24-hour or 7-day windows and behavioural features based on user profiles, velocity metrics, and geographic anomalies. Properly designed features provide 20%-40% performance enhancements over uncoded attributes (Yang et al., 2021; Kumar et al., 2025; Correa Bahnsen et al., 2016).
The minimum data requirements are three to six months of historical normal transaction data. The quality of data, unified merchant classification, and fingerprinting of devices are essential, and bad data quality severely worsens the performance of models (Ali et al., 2025; Kumar et al., 2025).

Strengths and limitations of current approaches

The existing machine learning-based methods have significant strengths. Hybrid Isolation Forest/LSTM-Autoencoder models are capable of production-scale performance with F1-scores over 0.90, and inference

times of less than 50ms (Priyanto et al., 2021; Tran et al., 2020). Ensemble approaches offer competitive performance that can be more readily interpreted. Deep learning models identify more complicated temporal and non-linear patterns than classical methods (Zapata-Cortes et al., 2024; Pumsirirat & Yan, 2018; Landauer et al., 2023; Moldovan & Iantovics, 2025).
Model performance is worse on imbalanced data in the real world than in controlled conditions, which indicates distributional overfitting (Budiansyah et al., 2025; Cavallaro et al., 2023). The streaming performance is not characterised, with most models being evaluated offline (Landauer et al., 2023; Al-Amri et al., 2021; Rafique et al., 2024). Deep learning models are not very transparent and thus cannot be used to triage incidents. Complex architectures are a processing cost burden which can be prohibitive on shared database server resources (Landauer et al., 2023; Al-Amri et al., 2021; Moldovan & Iantovics, 2025; Arshad et al., 2022; Ma et al., 2023). One-domain trained models, like the models that have been trained in the banking industry, are not generalizable to other domains like e-commerce without being fine-tuned (Dreshaj et al., 2025; Landauer et al., 2023; L. Zhang et al., 2024).

CONCLUSION

Summary of findings

This systematic literature review synthesises evidence from more than 43 investigations on machine learning-based anomaly detection in large-scale database systems and transaction environments. The main results correspond to the four research questions:
RQ1 - Types of anomalies: Database anomalies include business-level problems (fraud, unauthorised access) and technical-level problems (performance degradation, log errors, data integrity violations). There are three types of anomalies: point, contextual, and collective anomalies, which need different detection techniques. The transaction log anomalies complete the signals of transactional data analysis.
RQ2 - Machine Learning Technology: Isolation Forest is highly effective in computational efficiency when it comes to screening transactions in real time. Generative autoencoders, as well as autoencoders, identify nonlinear, complicated fraud patterns. The LSTM networks are useful in simulating time-dependent relationships within transaction chains and log event streams. Graph Neural Networks detect relational anomalies of transaction networks. Multi-technique composite methods are a current best practice, and Isolation Forest and LSTM-Autoencoder yield better performance to deploy in production. Deep Reinforcement Learning is promising when it comes to adaptive detection in dynamic workloads.
RQ3 - Problems: The imbalance in classes becomes extreme and requires unsupervised methods and the selection of metrics. Live requirements need effective models or tiered designs. Regulatory compliance and operational trust depend on interpretability. Concept drift needs to be monitored continuously and retrained periodically.
RQ4 - Practical Implications: Multi-tier architectures tradeoff between latency and accuracy. Extensive feature engineering has a significant impact on detection. The deployment of production must be monitored, A/B tested, and have a fallback. The basic requirements are data quality and compliance with privacy regulations. Gaps in the research and future directions.

Research gaps and future directions

The review provides a number of tangible gaps and future directions. To start with, lack of open, realistic and labelled database log and transaction anomaly benchmarks is a critical impediment to reproducibility. Community initiatives to create and release this type of dataset with the help of synthetic data generation and differential privacy are pressing (Dreshaj et al., 2025; Landauer et al., 2023).
Second, streaming and incremental machine learning designs should be designed and tested with real-time database constraints, and the detection latency, throughput effects, and computing overhead should be explicitly reported (Landauer et al., 2023; Al-Amri et al., 2021; Rafique et al., 2024; Cavallaro et al., 2023).
Third, explainable AI methods, such as SHAP, LIME, causal analysis, and attention-based mechanisms, must be logically implemented and tested in the database anomaly detection setting to allow generating actionable alerts (Ribeiro et al., 2016; Al-Amri et al., 2021; Moldovan & Iantovics, 2025).

Fourth, powerful domain-adaptive anomaly detection could be achieved with base models pre-trained on large corpora of transactions or logs (similar to BERT in the case of natural language processing). Other high-value future directions include GNNs to detect transaction network fraud rings, federated learning to detect privacy-preserving across organisations, and AutoML to democratize deployment (Ma et al., 2023; Motie & Raahemi, 2024; Reddy et al., 2025).
Lastly, domain-informed modelling that considers database-specific information like ACID properties, referential integrity constraints, workload periodicity, and access control policies must be explored to enhance the effectiveness of detection as well as minimise rates of false-positives in multi-tenant environments with complexities (Dreshaj et al., 2025; Wu et al., 2024).

Critical evaluation and actionable recommendations

In addition to the gap identification, critical analysis shows that although individual methods are promising, comparisons between them on standardised scales remain lacking. The field would be enriched with:
Reproducibility criteria: Hyper parameters, data preprocessing procedures and computation resources should be reported as mandatory, to allow fair comparison across studies.
Cross-domain validation: Trains models trained on banking data on e-commerce data to evaluate actual generalizability beyond the existing single-domain assessments.
Production deployment studies: Leaving offline experiments behind to publish actual deployment experience, such as integration costs, maintenance overhead, and operational issues.
Theoretical underpinnings: Coming up with formal structures of the reasons behind some architectures being better suited to some types of anomalies than others, rather than just trial-and-error.
These gaps are not merely technical challenges, but ones that are more concerned with the primordial questions of how machine learning can be reliably implemented in the mission-critical database environments.

Limitations of this review

There are a number of limitations to this review. English-language publications only were considered, and this could create language bias. Underrepresentation of negative results may occur due to publication bias. Even though six major databases were scanned, papers in specialised journals might not have been found. The review is done in a qualitative synthesis; no quantitative meta-analysis was carried out because the data, measurements and experimental procedures were heterogeneous. The imminent changes in machine learning require regular revision of these results.

Conclusion

In huge database systems and transaction platforms, machine learning has fundamentally revolutionised anomaly detection. The area has evolved beyond the rudimentary use of rules to monitor to advanced hybrid solutions that integrate a variety of machine learning methods. Isolation Forest offers a real-time, efficient baseline of detection. Deep learning, especially the LSTM-based autoencoders, learns intricate time-dependent dependencies and non-linear fraud patterns. Graph Neural Networks can be used in detecting relational structures. The trend toward multi-tier, hybrid solutions indicates that a more realistic understanding has emerged: no single technique can be the most effective for all types of anomalies and under all operating conditions.
Scalability, instant processing, explainability, and data standardisation issues are left critical. The gaps will need to be addressed through benchmark development, streaming architecture, explainable AI integration, and domain-aware modelling to make the full potential of machine learning-driven anomaly detection in production database settings a reality. The history of the field of work, from its separate methods to more integrated hybrid systems with operational rigour, has been toward an emerging discipline coming to a unified solution to one of the most significant problems of contemporary data infrastructure.

REFERENCES

1. Al-Amri, R., Murugesan, R. K., Man, M., Abdulateef, A. F., Al-Sharafi, M. A., & Alkahtani, A. A. (2021). A review of machine learning and deep learning techniques for anomaly detection in iot data. In Applied Sciences (Switzerland) (Vol. 11, Number 12). MDPI AG. https://doi.org/10.3390/app11125320
2. Aldweesh, A., Derhab, A., & Emam, A. Z. (2020). Deep learning approaches for anomaly-based intrusion detection systems: A survey, taxonomy, and open issues. Knowledge-Based Systems, 189, 105124. https://doi.org/10.1016/j.knosys.2019.105124
3. Ali, S., Boufaied, C., Bianculli, D., Branco, P., & Briand, L. (2025). A comprehensive study of machine learning techniques for log-based anomaly detection. Empirical Software Engineering, 30(5), 129. https://doi.org/10.1007/s10664-025-10669-3
4. Arshad, K., Ali, R. F., Muneer, A., Aziz, I. A., Naseer, S., Khan, N. S., & Taib, S. M. (2022). Deep Reinforcement Learning for Anomaly Detection: A Systematic Review. In IEEE Access (Vol. 10, pp. 124017–124035). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ACCESS.2022.3224023
5. Budiansyah, A., Zulfan, Z., Nizamuddin, N., Candra, R. A., Ilham, D. N., & Nazaruddin, N. (2025). The Effectiveness of Machine Learning Techniques in Anomaly Detection for Cyberattack Prevention: Systematic Literature Review 2020-2025. Brilliance: Research of Artificial Intelligence, 5(1), 259–271. https://doi.org/10.47709/brilliance.v5i1.6124
6. 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, 317–331. https://doi.org/10.1016/j.ins.2019.05.042
7. Cavallaro, C., Cutello, V., Pavone, M., & Zito, F. (2023). Discovering anomalies in big data: a review focused on the application of metaheuristics and machine learning techniques. In Frontiers in Big Data (Vol. 6). Frontiers Media SA. https://doi.org/10.3389/fdata.2023.1179625
8. Corli, S., Moro, L., Dragoni, D., Dispenza, M., & Prati, E. (2025). Quantum machine learning algorithms for anomaly detection: A review. Future Generation Computer Systems, 166, 107632. https://doi.org/10.1016/j.future.2024.107632
9. Correa Bahnsen, A., Aouada, D., Stojanovic, A., & Ottersten, B. (2016). Feature engineering strategies for credit card fraud detection. Expert Systems with Applications, 51, 134–142. https://doi.org/10.1016/j.eswa.2015.12.030
10. Dal Pozzolo, A., Caelen, O., Johnson, R. A., & Bontempi, G. (n.d.). Calibrating Probability with Undersampling for Unbalanced Classification.
11. Desai, A., Kosse, A., & Sharples, J. (2025). Finding a needle in a haystack: A machine learning framework for anomaly detection in payment systems. Journal of Finance and Data Science, 11. https://doi.org/10.1016/j.jfds.2025.100163
12. Dreshaj, A., Hamiti, M., Hasani, Z., Besimi, N., & Ajdari, J. (2025). Systematic Literature Review on Automatic Anomaly Detection Based on Database Logs. 2025 MIPRO 48th ICT and Electronics Convention, 1933–1937. https://doi.org/10.1109/MIPRO65660.2025.11132041
13. Du, M., Li, F., Zheng, G., & Srikumar, V. (2017). DeepLog: Anomaly detection and diagnosis from system logs through deep learning. Proceedings of the ACM Conference on Computer and Communications Security, 1285–1298. https://doi.org/10.1145/3133956.3134015
14. Fernando, T., Gammulle, H., Denman, S., Sridharan, S., & Fookes, C. (2022). Deep Learning for Medical Anomaly Detection A Survey. ACM Computing Surveys, 54(7). https://doi.org/10.1145/3464423
15. Guo, H., Yuan, S., & Wu, X. (2021). LogBERT: Log Anomaly Detection via BERT. 2021 International Joint Conference on Neural Networks (IJCNN), 1–8. https://doi.org/10.1109/IJCNN52387.2021.9534113
16. Hariri, S., Kind, M. C., & Brunner, R. J. (2021). Extended Isolation Forest. IEEE Transactions on Knowledge and Data Engineering, 33(4), 1479–1489. https://doi.org/10.1109/TKDE.2019.2947676

17. Hu, X., Xie, C., Fan, Z., Duan, Q., Zhang, D., Jiang, L., Wei, X., Hong, D., Li, G., Zeng, X., Chen, W., Wu, D., & Chanussot, J. (2022). Hyperspectral Anomaly Detection Using Deep Learning: A Review. In Remote Sensing (Vol. 14, Number 9). MDPI. https://doi.org/10.3390/rs14091973
18. Huang, S., Liu, Y., Fung, C., He, R., Zhao, Y., Yang, H., & Luan, Z. (2020). HitAnomaly: Hierarchical Transformers for Anomaly Detection in System Log. IEEE Transactions on Network and Service Management, 17(4), 2064–2076. https://doi.org/10.1109/TNSM.2020.3034647
19. Jurgovsky, J., Granitzer, M., Ziegler, K., Calabretto, S., Portier, P.-E., He-Guelton, L., & Caelen, O. (2018). Sequence classification for credit-card fraud detection. Expert Systems with Applications, 100, 234–245. https://doi.org/10.1016/j.eswa.2018.01.037
20. Kitchenham, B., Pearl Brereton, O., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering – A systematic literature review. Information and Software Technology, 51(1), 7–15. https://doi.org/10.1016/j.infsof.2008.09.009
21. Kumar, A., Kumar, A., Raja, R., Dewangan, A. K., Kumar, M., Soni, A., Agarwal, D., & Saudagar, A. K. J. (2025). Revolutionising anomaly detection: a hybrid framework for anomaly detection integrating isolation forest, autoencoder, and Conv. LSTM. Knowledge and Information Systems, 67(12), 11903–11953. https://doi.org/10.1007/s10115-025-02580-6
22. Landauer, M., Onder, S., Skopik, F., & Wurzenberger, M. (2023). Deep learning for anomaly detection in log data: A survey. Machine Learning with Applications, 12, 100470. https://doi.org/10.1016/j.mlwa.2023.100470
23. Liu, F. T., Ting, K. M., & Zhou, Z.-H. (2008). Isolation Forest. 2008 Eighth IEEE International Conference on Data Mining, 413–422. https://doi.org/10.1109/ICDM.2008.17
24. Ma, X., Wu, J., Xue, S., Yang, J., Zhou, C., Sheng, Q. Z., Xiong, H., & Akoglu, L. (2023). A Comprehensive Survey on Graph Anomaly Detection With Deep Learning. IEEE Transactions on Knowledge and Data Engineering, 35(12), 12012–12038. https://doi.org/10.1109/TKDE.2021.3118815
25. Meng, W., Liu, Y., Zhu, Y., Zhang, S., Pei, D., Liu, Y., Chen, Y., Zhang, R., Tao, S., Sun, P., & Zhou, R. (n.d.). LogAnomaly: Unsupervised Detection of Sequential and Quantitative Anomalies in Unstructured Logs.
26. Moldovan, S. C., & Iantovics, L. B. (2025). Review on Information Fusion-Based Data Mining for Improving Complex Anomaly Detection. In Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery (Vol. 15, Number 2). John Wiley and Sons Inc. https://doi.org/10.1002/widm.70017
27. Motie, S., & Raahemi, B. (2024). Financial fraud detection using graph neural networks: A systematic review. Expert Systems with Applications, 240, 122156. https://doi.org/10.1016/j.eswa.2023.122156
28. Nassif, A. B., Talib, M. A., Nasir, Q., & Dakalbab, F. M. (2021). Machine Learning for Anomaly Detection: A Systematic Review. In IEEE Access (Vol. 9, pp. 78658–78700). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ACCESS.2021.3083060
29. Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. In BMJ (Vol. 372). BMJ Publishing Group. https://doi.org/10.1136/bmj.n71
30. Priyanto, C. Y., Hendry, & Purnomo, H. D. (2021). Combination of Isolation Forest and LSTM Autoencoder for Anomaly Detection. 2021 2nd International Conference on Innovative and Creative Information Technology (ICITech), 35–38. https://doi.org/10.1109/ICITech50181.2021.9590143

31. Pumsirirat, A., & Yan, L. (2018). Credit Card Fraud Detection using Deep Learning based on Auto-Encoder and Restricted Boltzmann Machine. In IJACSA) International Journal of Advanced Computer Science and Applications (Vol. 9, Number 1). www.ijacsa.thesai.org
32. Rafique, S. H., Abdallah, A., Musa, N. S., & Murugan, T. (2024). Machine Learning and Deep Learning Techniques for Internet of Things Network Anomaly Detection—Current Research Trends. Sensors, 24(6), 1968. https://doi.org/10.3390/s24061968
33. Reddy, C., Prabhagaran, S., & Vaid, A. (2025). Adaptive Anomaly Detection in Database Transactions: Bridging Security Gaps with Reinforcement Learning. European Journal of Artificial Intelligence and Machine Learning, 4(2), 8–14. https://doi.org/10.24018/ejai.2025.4.2.53
34. Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). 'Why should i trust you?' Explaining the predictions of any classifier. Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 13-17-August-2016, 1135–1144. https://doi.org/10.1145/2939672.2939778
35. Saraswathi, S., & Selvakumar, S. (2025). Enhanced Anomaly Detection in Wireless Sensor Networks Using Isolation Forest, LSTM, and LSTM Autoencode. 2025 3rd International Conference on Inventive Computing and Informatics (ICICI), 1528–1534. https://doi.org/10.1109/ICICI65870.2025.11069875
36. Tahir, M., Abdullah, A., Izura Udzir, N., & Azhar Kasmiran, K. (2025). A systematic review of machine learning and deep learning techniques for anomaly detection in data mining. International Journal of Computers and Applications, 47(2), 169–187. https://doi.org/10.1080/1206212X.2025.2449999
37. Tran, P. H., Heuchenne, C., & Thomassey, S. (2020). An anomaly detection approach based on the combination of LSTM autoencoder and isolation forest for multivariate time series data. Developments of Artificial Intelligence Technologies in Computation and Robotics, 589–596. https://doi.org/10.1142/9789811223334_0071
38. Wu, Y., Sicard, B., & Gadsden, S. A. (2024). Physics-informed machine learning: A comprehensive review on applications in anomaly detection and condition monitoring. In Expert Systems with Applications (Vol. 255). Elsevier Ltd. https://doi.org/10.1016/j.eswa.2024.124678
39. Yang, L., Chen, J., Wang, Z., Wang, W., Jiang, J., Dong, X., & Zhang, W. (2021). Semi-Supervised Log-Based Anomaly Detection via Probabilistic Label Estimation. 2021 IEEE/ACM 43rd International Conference on Software Engineering (ICSE), 1448–1460. https://doi.org/10.1109/ICSE43902.2021.00130
40. Yoon, D. Y., Niu, N., & Mozafari, B. (2016). DBSherlock: A performance diagnostic tool for transactional databases. Proceedings of the ACM SIGMOD International Conference on Management of Data, 26-June-2016, 1599–1614. https://doi.org/10.1145/2882903.2915218
41. Zapata-Cortes, O., Arango-Serna, M. D., Zapata-Cortes, J. A., & Restrepo-Carmona, J. A. (2024). Machine Learning Models and Applications for Early Detection. In Sensors (Vol. 24, Number 14). Multidisciplinary Digital Publishing Institute (MDPI). https://doi.org/10.3390/s24144678
42. Zhang, L., Jia, T., Jia, M., Li, Y., Yang, Y., & Wu, Z. (2024). Multivariate Log-based Anomaly Detection for Distributed Database. Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 4256–4267. https://doi.org/10.1145/3637528.3671725
43. Zhang, X., Xu, Y., Lin, Q., Qiao, B., Zhang, H., Dang, Y., Xie, C., Yang, X., Cheng, Q., Li, Z., Chen, J., He, X., Yao, R., Lou, J. G., Chintalapati, M., Shen, F., & Zhang, D. (2019). Robust log-based anomaly detection on unstable log data. ESEC/FSE 2019 - Proceedings of the 2019 27th ACM Joint Meeting European Software Engineering Conference and Symposium on the Foundations of Software Engineering, 807–817. https://doi.org/10.1145/3338906.3338931