

ANOMALY DETECTION IN TIME SERIES DATA: TRENDS, APPLICATIONS, AND RESEARCH GAPS

Mr. Rushi Raval^{1*}, Dr. Tejas Patalia²

^{1*}Research Scholar, Gujarat Technological University, Ahmedabad, rushiraval07@gmail.com

²Professor, V.V.P Engineering College, Rajkot, pataliatejas@rediffmail.com

Abstract

Time series anomaly detection plays a pivotal role in modern intelligent systems, enabling the identification of irregular patterns within sequential data across domains such as finance, healthcare, manufacturing, and cybersecurity. With the rapid proliferation of IoT devices and sensors generating massive real-time data streams, detecting anomalies has become increasingly critical yet complex. This review comprehensively examines recent advancements in anomaly detection techniques, encompassing statistical models, traditional machine learning algorithms, and emerging deep learning architectures. It contrasts classical methods such as ARIMA and STL decomposition with advanced models like RNNs, LSTMs, autoencoders, and transformer-based frameworks, emphasizing their comparative strengths in scalability, interpretability, and performance. Additionally, the review highlights cross-domain applications, including fraud detection, predictive maintenance, disease surveillance, and environmental monitoring, demonstrating the widespread utility of anomaly detection systems. Emerging trends such as explainable AI, federated and edge learning, and privacy-preserving frameworks are explored as key enablers of trustworthy and adaptive solutions. The study also identifies research gaps, including the scarcity of benchmark datasets, challenges in handling high-dimensional and multivariate data, and issues of generalizability across domains. By consolidating existing literature, this review proposes a taxonomy of techniques and outlines future directions for developing transparent, efficient, and domain-agnostic anomaly detection frameworks.

Keywords: Time series anomaly detection, Deep learning, Explainable AI, Internet of Things (IoT), Predictive analytics

1. Introduction

Time-series anomaly detection has become indispensable for intelligent systems that monitor, predict, and manage real-world processes in finance, healthcare, manufacturing, and cybersecurity. The unprecedented growth of Internet of Things (IoT) sensors, connected infrastructures, and streaming data has intensified the need for algorithms capable of identifying irregular patterns in real time (Nizam et al., 2022). Anomalies, defined as deviations from expected temporal behavior, often signal potential faults, security breaches, or critical transitions in dynamic environments. Traditional statistical models such as ARIMA, Holt–Winters, and Kalman filters provided early foundations for anomaly detection but are constrained by their linear assumptions and limited scalability (Desani & Chittibala, 2021). With the increasing complexity of data streams, adaptive machine learning and deep learning techniques have transformed the field by enabling automatic feature extraction and contextual understanding. These innovations have redefined anomaly detection from a reactive process to a proactive analytical discipline that enhances situational awareness and decision-making across domains (Sharif et al., 2022).

Over the past few years, anomaly detection research has diversified across statistical, machine learning, and deep learning paradigms, evolving into one of the most interdisciplinary areas of temporal data science. Traditional algorithms, though interpretable, struggle to handle high-dimensional and nonstationary time-series data. Deep architectures such as recurrent neural networks (RNNs), long short-term memory (LSTM) networks, and autoencoders have demonstrated superior ability to capture nonlinear temporal dependencies, significantly improving accuracy and robustness

in dynamic environments (Karadayi et al., 2020). Industrial IoT applications have particularly benefited from these methods, as multivariate time-series data from sensors often exhibit complex correlations that require advanced representational learning to detect subtle and contextual anomalies (Nizam et al., 2022).

Recent surveys emphasize the growing role of explainable artificial intelligence (XAI) in anomaly detection, highlighting the necessity of transparency and interpretability in model outputs (Li et al., 2023). Black-box models, while powerful, can limit trust and practical deployment in critical sectors such as cybersecurity and healthcare. Explainable approaches use attention visualization, attribution maps, and saliency analysis to elucidate why specific patterns are flagged as anomalous, improving both model reliability and human oversight. Parallel to this, research in federated and privacy-preserving learning has introduced collaborative anomaly detection frameworks that maintain data confidentiality while leveraging distributed computational resources (Kumar et al., 2021). These developments are particularly valuable for large-scale industrial systems and cross-institutional networks where data sharing is restricted.

Complementary works have examined hybrid models, combining forecasting and reconstruction paradigms to improve adaptability to temporal drift and nonstationary data. Sparse neural networks and unsupervised representation learning have further reduced dependency on labeled data, making these systems applicable to emerging fields such as biomedical monitoring and fault detection (Gugulothu et al., 2018; Pereira & Silveira, 2019). Moreover, reviews such as Shaukat et al. (2021) underscore persistent gaps in benchmark consistency, scalability, and domain transferability. Collectively, the literature demonstrates significant progress toward real-time, explainable, and privacy-aware anomaly detection systems, but also reveals ongoing challenges in standardization, interpretability, and cross-domain generalization (Ali, 2024).

This review synthesizes the evolution of time-series anomaly detection across statistical, machine learning, and deep learning domains, emphasizing the comparative strengths and limitations of each paradigm. It covers forecasting, reconstruction, density-based, and hybrid approaches across univariate and multivariate time-series contexts. The discussion extends to explainable, federated, and privacy-preserving frameworks that are reshaping anomaly detection for modern intelligent systems. Through a taxonomy of supervision, paradigm, anomaly type, and deployment context, this review provides an integrated framework for understanding methodological progress and practical applications.

The objectives of this review are to establish a unified taxonomy of time-series anomaly detection methods, enabling systematic comparison across statistical, machine learning, and deep learning paradigms. It further aims to explore the applicability of these techniques across major industrial and research domains, including finance, healthcare, manufacturing, and cybersecurity. Additionally, the review identifies existing limitations and emerging directions related to interpretability, scalability, and domain adaptation. By fulfilling these objectives, the study seeks to guide future research toward transparent, adaptive, and domain-agnostic frameworks that enhance reliability, efficiency, and trust in intelligent, data-driven anomaly detection systems.

2. Foundations of Time-Series Anomaly Detection

2.1 Characteristics of Time-Series Data

Time-series data represent sequential observations collected over time, characterized by temporal dependencies, seasonality, and evolving trends. Each observation is often correlated with preceding and succeeding values, which introduces autocorrelation and makes standard statistical assumptions of independence invalid. These dependencies mean that effective anomaly detection must distinguish between normal temporal variation and true irregularities (Zamanzadeh Darban et al., 2024). Seasonality captures recurring patterns such as daily or annual cycles, whereas trends reflect long-term systematic changes. Furthermore, time-series data can be univariate, containing a single attribute, or multivariate, involving several correlated dimensions that evolve together, such as multiple sensors

in industrial systems or multichannel physiological signals in healthcare. Real-world series are typically nonstationary as their statistical properties shift over time due to context, sensor aging, or system evolution. Moreover, the presence of noise and missing data can obscure meaningful patterns, making model robustness essential (Schmidl et al., 2022). Consequently, reliable anomaly detection requires methods capable of handling complex temporal correlations, high-dimensional interdependencies, and nonstationary noise conditions.

2.2 Types of Anomalies

Anomalies in time-series data generally fall into three categories: point, contextual, and collective. A point anomaly refers to a single observation that deviates significantly from the expected range of normal behavior. Contextual anomalies depend on the surrounding temporal context. A data point may appear normal overall but abnormal within a specific local window. For instance, an unusually high temperature reading might be anomalous in winter but not in summer (Hundman et al., 2018). Collective anomalies involve a set of consecutive observations that appear normal individually but collectively exhibit a pattern inconsistent with historical norms, such as gradually increasing vibration amplitudes before a machine failure or subtle drifts in satellite telemetry before malfunction. Each anomaly type requires distinct modeling approaches: statistical thresholds and forecasting models often handle point anomalies, while sequence-based and deep temporal models capture contextual or collective patterns (Zamanzadeh Darban et al., 2024). Recognizing these distinctions is essential when mapping problem types to appropriate algorithms and when evaluating their performance across domains such as finance, healthcare, manufacturing, and environmental systems.

2.3 Problem Formulation and Evaluation Metrics

Formally, a time series can be represented as:

$$X = \{x_1, x_2, \dots, x_T\}$$

where x_t denotes the value observed at time step t . The objective of anomaly detection is to learn a scoring function

$$f: X \rightarrow S = \{s_1, s_2, \dots, s_T\}$$

that produces anomaly scores s_t for each timestamp. A binary decision function then determines whether a point is anomalous using a predefined threshold θ :

$$y_t = \begin{cases} 1, & \text{if } s_t > \theta \\ 0, & \text{otherwise} \end{cases}$$

Performance evaluation depends on the application's tolerance for false alarms versus missed detections. Classical measures such as Precision, Recall, and F1-score quantify the accuracy of detected anomalies, while ROC-AUC and PR-AUC provide threshold-independent assessments of detection performance. For sequential data, range-based metrics extend these concepts to contiguous anomaly intervals rather than isolated points, improving evaluation realism in temporal contexts (Lee et al., 2018). Recent studies have also emphasized latency-aware metrics, which penalize delayed detections to account for the importance of early alerts in high-stakes applications such as spacecraft monitoring and industrial automation (Sørbø & Ruocco, 2024; Paparrizos et al., 2025). Data preprocessing plays a critical role in improving model sensitivity and stability. Normalization ensures amplitude comparability across signals, decomposition techniques such as STL isolate trend and seasonal components, and windowing converts continuous streams into fixed-length overlapping segments suitable for model input. These preprocessing operations mitigate the effects of periodicity and nonstationarity, reducing bias and enhancing the precision of both statistical and deep learning approaches (Schmidl et al., 2022). The principal anomaly types and representative cross-domain examples are summarized in Table 1.

Table 1. Types of anomalies in time-series data and domain examples

Anomaly Type	Definition	Example Domain Instance
--------------	------------	-------------------------

Point anomaly	A single observation that deviates sharply from normal expectations	Sudden surge in transaction value (finance)
Contextual anomaly	A point anomalous only within a specific temporal context	Elevated heart rate during rest (healthcare)
Collective anomaly	A sequence of values jointly abnormal though individually plausible	Progressive vibration rise before machine failure (manufacturing)

3. Classical Statistical and Machine Learning Approaches

3.1 Statistical Methods

Classical statistical models provide the earliest formal foundations for time-series anomaly detection. Among these, the Autoregressive Integrated Moving Average (ARIMA) model remains central. It models temporal dependence through autoregressive and moving-average terms and identifies anomalies when residuals deviate significantly from their expected variance (Sun et al., 2024). The model's transparency and strong theoretical basis make it widely adopted in forecasting-driven domains, although its linearity and assumption of stationarity restrict its performance for dynamic or highly nonlinear data streams (Hoeltgebaum et al., 2021).

To address seasonality and trend, the Holt–Winters exponential-smoothing approach introduces adaptive parameters for level, trend, and seasonal components. It performs particularly well in univariate sensor and energy data with cyclical behavior (Sun et al., 2024). Likewise, STL decomposition (Seasonal–Trend decomposition using Loess) isolates deterministic components from residual irregularities, allowing anomalies to be detected directly in the residual series (Braei & Wagner, 2020).

The Kalman filter, another statistical cornerstone, treats observations as noisy estimates of latent system states evolving. By recursively estimating these hidden states and associated covariance, it can flag anomalies when prediction innovations exceed probabilistic thresholds (Puder et al., 2024). Change-point detection methods such as Cumulative Sum (CUSUM) and Exponentially Weighted Moving Average (EWMA) complement these models by monitoring structural shifts in a signal's mean or variance (Truong et al., 2020). CUSUM rapidly detects abrupt deviations, while EWMA sensitively captures gradual drifts and persistent bias changes.

The strengths of these techniques lie in their interpretability, analytical tractability, and low data requirements. Every parameter has a clear physical meaning, such as trend, variance, or smoothing constant, which enables transparent decision making (Braei & Wagner, 2020). However, their limitations become evident in nonstationary, high-dimensional, and nonlinear contexts. In real-time IoT and industrial cyber-physical systems, statistical models often misclassify noise bursts or dynamic shifts as anomalies because they lack adaptive capacity (Hoeltgebaum et al., 2021). Recent studies have therefore explored hybrid statistical–machine-learning frameworks, where ARIMA or Kalman residuals feed into learning-based detectors, improving detection robustness while retaining interpretability (Hao et al., 2021; Puder et al., 2024). Despite their simplicity, classical models remain indispensable for benchmarking and for applications requiring transparency and fast interpretability (Mandaleeka & Bonthu, 2025). Figure 1 summarizes how these foundations evolve toward more data-driven methods.

3.2 Traditional Machine Learning Methods

Traditional machine-learning (ML) algorithms extend anomaly detection to nonlinear and multivariate domains by inferring patterns directly from data. Clustering-based methods such as k-means and DBSCAN classify subsequences into dense regions, designating sparsely located points as anomalies (Braei & Wagner, 2020). K-means performs efficiently but requires a predefined cluster count and assumes spherical clusters, whereas DBSCAN detects arbitrarily shaped clusters yet is sensitive to the choice of neighborhood parameters (Inuwa & Das, 2024). Distance-based techniques compute dissimilarities among subsequences using Euclidean or Dynamic Time Warping (DTW)

measures; those with the largest average distance are treated as outliers. While intuitive, such approaches scale poorly for long sequences or streaming contexts (Meng et al., 2023). The Isolation Forest (IF) algorithm isolates anomalies through random recursive partitioning. Anomalous points require fewer partitions to be isolated. This unsupervised ensemble technique has inspired enhanced versions such as Extended IF and Deep IF for nonlinear, high-dimensional time series (Xu et al., 2023). Similarly, One-Class Support Vector Machine (OC-SVM) constructs a boundary that encloses normal data, identifying points outside the boundary as anomalous, while the Local Outlier Factor (LOF) computes the relative local density of each instance, flagging regions with sparse density (Ruff et al., 2018).

The performance of these ML methods depends heavily on feature construction and representation. Statistical descriptors (mean, variance, autocorrelation), frequency-domain components, or learned embeddings transform time series into fixed-length vectors suitable for ML models (Hamon et al., 2025). Representation quality directly impacts detection accuracy. Recent reviews highlight that unsupervised feature learning enhances generalization across domains while reducing manual engineering (Meng et al., 2023).

Compared with statistical models, ML algorithms exhibit greater flexibility, scalability, and nonlinearity handling, yet their interpretability and stability are often limited. Parameters such as contamination rate, kernel width, or neighborhood radius strongly influence outcomes, requiring extensive tuning (Inuwa & Das, 2024). Hybrid models, combining statistical forecasting with ML scoring or fusing residual analysis with clustering, offer a balanced approach, yielding improved robustness and domain transferability (Hao et al., 2021; Mandaleeka & Bonthu, 2025). Despite the rise of deep learning, classical ML methods remain competitive due to their explainability, efficiency, and adaptability across IoT, finance, and industrial control scenarios. The methodological evolution of time-series anomaly detection has advanced from classical statistical baselines toward scalable and generalizable deep architectures (Figure 1).

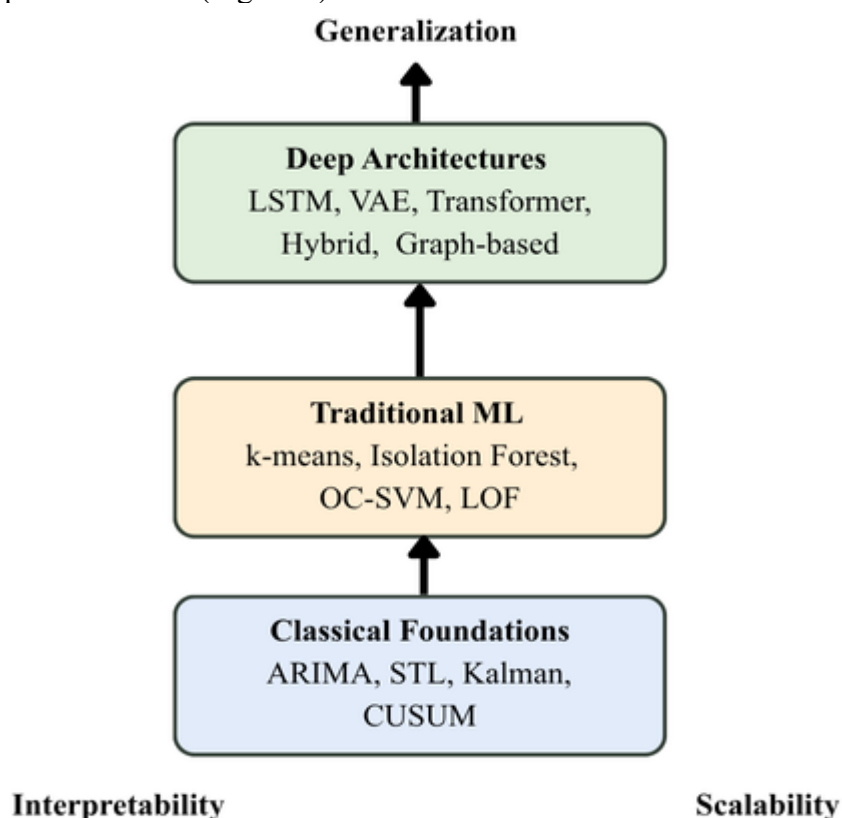


Figure 1: Evolution of time-series anomaly detection methods from classical statistical models to scalable and generalizable deep learning architectures

Table 2 contrasts these approaches and their trade-offs in interpretability, scalability, and performance.

Table 2. Comparison of classical statistical and traditional ML methods for time-series anomaly detection

Method Class	Representative Techniques	Strengths	Limitations
Statistical	ARIMA, STL, Holt-Winters, Kalman, CUSUM, EWMA	Transparent, interpretable, computationally efficient	Limited to linear, stationary data; poor scalability
Traditional ML	k-means, DBSCAN, IF, OC-SVM, LOF	Flexible, nonlinear modeling, handles multivariate data	Requires feature engineering, sensitive to hyperparameters, less interpretable

4. Deep Learning Architectures for Anomaly Detection

4.1 Forecasting-Based Models

Forecasting-based deep architectures predict future time points and quantify anomalies through error analysis. Recurrent Neural Networks (RNNs) and their variants, such as Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) networks, are commonly used to model temporal dependencies because they retain long-term context across sequences. When the difference between predicted and observed values exceeds an adaptive threshold, it signals an anomaly (Choi et al., 2021). Temporal Convolutional Networks (TCNs) extend this concept using dilated causal convolutions, capturing long-range dependencies without recurrence and enabling parallel processing (Iqbal & Amin, 2024). Forecasting-based models are widely applied in energy systems, finance, and industrial process monitoring, where stable cyclicity allows deviations to be detected via residual error (Huang et al., 2025). However, they are limited when regimes shift abruptly or anomalies occur collectively rather than as isolated points.

4.2 Reconstruction-Based Models

Reconstruction-based models learn to reproduce normal patterns so that abnormal inputs reconstruct poorly. Autoencoders (AEs) encode input sequences into latent representations and then reconstruct them; high reconstruction error denotes abnormality. Variational Autoencoders (VAEs) introduce probabilistic latent variables, measuring deviation via both reconstruction loss and log-likelihood (Li et al., 2020). LSTM-Autoencoders (LSTM-AEs) combine recurrent units within the encoder–decoder to preserve sequential continuity across sliding windows (Choi et al., 2021). Generative Adversarial Networks (GANs) such as TadGAN apply adversarial learning, where the generator reconstructs plausible sequences and the discriminator identifies unrealistic ones; anomaly scores combine reconstruction error and discriminator feedback (Geiger et al., 2020). These models are especially valuable in unsupervised contexts because they learn from normal data alone. Their strength lies in modeling nonlinear structures in high-dimensional data, though they require careful tuning and are sensitive to training instability and computational cost.

4.3 Transformer-Based and Hybrid Architectures

Transformer architectures have redefined time-series anomaly detection by leveraging self-attention mechanisms to model long-range dependencies without recurrence. Attention weights allow the network to focus selectively on relevant timestamps, overcoming gradient decay and capturing contextual patterns effectively (Kim et al., 2023). In multivariate settings, graph-based transformers model correlations between sensors or variables using graph neural network layers integrated with attention modules, enhancing both accuracy and explainability (Jin et al., 2024). Hybrid designs merge forecasting and reconstruction paradigms, for example by embedding transformer encoders with

recurrent or convolutional decoders, or fusing attention blocks with autoencoders to learn dual representations (Huang et al., 2025). These systems achieve strong scalability and data-efficiency across domains ranging from IoT to climate analytics. Nevertheless, transformers introduce new trade-offs: they demand heavy computational resources, are prone to overfitting on small datasets, and remain challenging to interpret despite visualization advances in attention mapping (Choi et al., 2021).

4.4 Comparative Evaluation

Comparative evaluations show distinct performance characteristics among forecasting, reconstruction, and attention-driven deep models. Benchmarks such as the Numenta Anomaly Benchmark (NAB), Yahoo S5, and UCR Time Series Archive provide standardized testbeds, with metrics including Precision, Recall, F1-Score, ROC-AUC, and PR-AUC used for consistent comparison (Huang et al., 2025). Forecasting-based models perform best for point anomalies in stable periodic data, while reconstruction-based approaches excel for contextual or collective anomalies where patterns deviate structurally (Li et al., 2020). Transformer and hybrid frameworks outperform others on multivariate datasets that require learning cross-variable dependencies (Kim et al., 2023). Despite their accuracy, deep models still face challenges of interpretability, data imbalance, and computational cost. Emerging solutions, such as graph-based reasoning, attention visualization, and self-supervised pretraining, aim to improve transparency and adaptability (Jin et al., 2024). Table 3 summarizes representative deep architectures, detailing input types, learning modes, and evaluation criteria, thereby illustrating the evolution toward domain-agnostic, scalable frameworks for time-series anomaly detection.

Table 3. Summary of deep learning-based approaches

Approach	Input Type	Architecture	Supervision	Primary Metrics
Forecasting (RNN, LSTM, GRU, TCN)	Sequence windows	Recurrent / Convolutional	Self-supervised	RMSE, MAE, Z-score error
Reconstruction (AE, VAE, LSTM-AE)	Full sequence	Encoder–Decoder	Unsupervised	Reconstruction error, log-likelihood
GAN-based (TadGAN)	Subsequence windows	Generator + Discriminator	Unsupervised	Combined reconstruction + discriminator loss
Transformer / Hybrid	Multivariate series	Attention / Graph / Hybrid	Self- or semi-supervised	F1, PR-AUC, ROC-AUC, Latency-aware metrics

5. Cross-Domain Applications

5.1 Finance: Fraud Detection and Risk Analytics

In the financial domain, anomaly detection is a frontline defense for identifying irregular patterns in high-velocity transaction streams such as payments, trading activities, and insurance claims. Recent approaches fuse graph neural networks (GNNs) with sequential anomaly detection to model relational dependencies across accounts and detect sophisticated, coordinated fraud in real time (Rasul et al., 2024). Cost-sensitive learning further reduces false-alarm rates by penalizing unnecessary interventions while maintaining detection sensitivity (Zhu et al., 2021). Hierarchical embedding mechanisms and attention-based temporal modeling enable early detection of abnormal financial behaviors, such as abrupt increases in transaction volumes or deviations in transfer patterns. These systems move beyond static rule engines toward adaptive and explainable fraud analytics frameworks that align model interpretability with regulatory transparency in modern financial auditing.

5.2 Healthcare: Disease Surveillance and Patient Monitoring

In healthcare, anomaly detection drives predictive and preventive intelligence by continuously monitoring physiological signals and detecting early pathological deviations. Real-time ECG, heart-rate, and oxygen-saturation signals are modeled as temporal sequences in which subtle drifts may precede clinical deterioration. Transformer-based unsupervised models now achieve high accuracy in detecting cardiac or respiratory anomalies without labeled data (Alamr & Artoli, 2023). Complementary research has integrated hierarchical federated learning and digital-twin architectures, where hospital-edge nodes train local models on patient data while contributing to global anomaly detection without breaching privacy (Gupta et al., 2021). These intelligent, privacy-preserving systems balance sensitivity and specificity, reducing alarm fatigue and enabling continuous, adaptive surveillance across distributed healthcare environments.

5.3 Manufacturing: Predictive Maintenance and Fault Detection

Industrial anomaly detection underpins predictive maintenance in cyber-physical manufacturing systems, where continuous sensing captures vibration, current, temperature, and torque data from robots and machinery. Deep hybrid architectures that combine convolutional autoencoders with physical process models identify degradation patterns before visible failures occur (Wang et al., 2024). By embedding physical constraints into digital-twin frameworks, these systems provide both accuracy and causal interpretability, allowing engineers to trace fault origins. In robotic assembly lines, time-series anomalies trigger proactive interventions that minimize downtime and optimize maintenance schedules. This integration of physics-informed learning with data-driven adaptation exemplifies how industrial anomaly detection contributes to reliable, explainable, and self-correcting smart-factory operations.

5.4 Cybersecurity and Network Monitoring

Cybersecurity environments generate massive, rapidly evolving time-series logs encompassing packet traffic, access events, and system metrics. Deep graph-temporal fusion (GTF) networks have emerged as powerful tools for intrusion detection, capturing both temporal and relational dependencies within in-vehicle CAN or enterprise networks (Yang et al., 2024). Such models distinguish benign irregularities from coordinated malicious behavior with high precision. Hybrid edge-cloud frameworks deploy lightweight models close to network gateways for rapid threat localization while central servers handle global anomaly correlation. As attackers increasingly employ adversarial techniques, adaptive and semi-supervised detectors that continuously recalibrate on new data streams ensure robust, real-time defense. Consequently, anomaly detection has evolved from passive monitoring to dynamic cyber-resilience, aligning with the adaptive, intelligent-system vision outlined in the abstract.

5.5 Environmental and Climate Monitoring

In environmental science, anomaly detection identifies extreme weather events, environmental hazards, and sensor faults within vast spatio-temporal datasets. Spatio-temporal autoencoders have been widely adopted to detect unusual climatic patterns, leveraging spatial correlation across atmospheric grids and temporal consistency across time (Tibau et al., 2021). Similarly, time-distributed CNN-LSTM architectures enhance flood-forecasting accuracy by modeling sequential hydrological interactions across rainfall, runoff, and soil-moisture indices (Malik et al., 2024). For smart-city management, unsupervised IoT anomaly detectors monitor environmental sensors to detect air-quality violations, infrastructure malfunctions, or urban heat-island effects in real time (Guo et al., 2020). These applications highlight how time-series anomaly detection enables sustainable and climate-resilient infrastructure through automated event detection and decision support for public safety.

6. Emerging Trends and Technological Enablers

6.1 Explainable Artificial Intelligence (XAI)

The interpretability of deep learning models has become crucial for anomaly detection in time-series data, especially in domains like healthcare, finance, and industrial monitoring. Explainable Artificial Intelligence (XAI) bridges the gap between model accuracy and human understanding by highlighting which time steps or input features influence anomaly decisions. Methods such as attention visualization, saliency mapping, and attribution analysis are increasingly used to interpret RNN, CNN, and transformer outputs (Rojat et al., 2021). Transformer-based models, while powerful for sequential analysis, are often opaque due to multi-head attention layers and latent temporal embeddings. Recent architectures, such as DeepECG-Net, integrate hybrid transformers with interpretable attention heads to detect and visualize cardiac anomalies in real time, achieving transparency without compromising diagnostic precision (Alghieth, 2025). Similarly, healthcare-focused reviews emphasize that integrating attention heatmaps and relevance propagation techniques enhances clinician trust in AI-driven anomaly detection (Yang et al., 2023). XAI also supports causal interpretability, where counterfactual perturbations reveal minimal changes that flip anomaly scores, thereby explaining why certain temporal patterns are flagged as abnormal. Nevertheless, key challenges persist in ensuring temporal consistency of explanations, maintaining computational efficiency in real-time systems, and generalizing interpretability across diverse multivariate datasets.

6.2 Federated and Edge Learning

The explosion of IoT and sensor networks has created vast volumes of distributed time-series data that cannot be centralized due to privacy, bandwidth, or latency constraints. Federated Learning (FL) addresses this by training models collaboratively across devices while retaining data locally. In the context of anomaly detection, FL improves cross-domain adaptability and privacy preservation but must contend with client heterogeneity and communication efficiency (Zhang et al., 2021). Recent frameworks such as PeFAD, a parameter-efficient federated architecture, optimize communication by fine-tuning a small subset of parameters at each client while maintaining model accuracy across industrial and healthcare data streams (Xu et al., 2024). Other studies employ reservoir state analysis to detect anomalies directly on edge devices, reducing synchronization costs and mitigating non-IID data distribution issues (Nogami, Tamura, & Tanaka, 2025). To ensure robust performance under privacy constraints, the FedSW-TSAD model leverages Sobolev–Wasserstein GANs to improve stability and adaptability in multivariate time-series detection across heterogeneous clients (Zhang X. et al., 2025). These innovations collectively enhance scalability, privacy, and communication efficiency, paving the way for real-time distributed anomaly detection at the edge.

6.3 Privacy-Preserving and Trustworthy Frameworks

As anomaly detection systems proliferate in sensitive environments, ensuring privacy, robustness, and trust has become an essential design priority. Privacy-preserving frameworks combine cryptographic aggregation, differential privacy, and lightweight encryption to protect both data and model parameters. For example, GuardianAI implements differential-privacy noise injection and secure aggregation to prevent gradient leakage in federated anomaly detection setups, ensuring end-to-end confidentiality during distributed training (Alabdulatif, 2025). In industrial IoT environments, privacy must coexist with computational efficiency; thus, recent frameworks introduce lightweight encrypted autoencoders and secure aggregation layers that maintain detection precision under limited hardware capacity (Chen et al., 2025). Complementary advances focus on human-in-the-loop calibration, where experts review low-confidence alerts to adjust thresholds dynamically, and on uncertainty quantification methods that model the confidence of each anomaly prediction. Explainable and trustworthy AI approaches, such as those combining XAI with anomaly forecasting, enhance interpretability and accountability by visualizing decision pathways and uncertainty distributions (Iqbal & Amin, 2025). Together, these frameworks enable responsible, transparent, and resilient

anomaly detection suitable for critical domains like healthcare, finance, and smart-industry infrastructures.

7. Research Gaps and Open Challenges

7.1 Lack of Standardized Benchmark Datasets

A significant barrier to progress in time-series anomaly detection research is the absence of standardized and widely accepted benchmark datasets. Many studies rely on proprietary or domain-specific datasets, making cross-comparison of models nearly impossible. Existing datasets often contain limited labeled anomalies, leading to biased evaluation and overfitting to narrow contexts. Synthetic datasets, while useful for controlled experiments, fail to capture the stochasticity, noise, and structural heterogeneity found in real-world systems. Moreover, anomaly annotations are frequently ambiguous or inconsistent, resulting in unreliable metrics. To advance the field, there is a growing demand for domain-agnostic, range-based, and drift-aware benchmark datasets that reflect real temporal irregularities, concept drift, and varying anomaly durations. Establishing open, curated repositories with transparent annotation protocols would significantly improve reproducibility and the generalization of anomaly detection models across domains.

7.2 High-Dimensional and Multivariate Complexity

As modern systems generate increasingly multivariate data streams ranging from industrial IoT sensors to multi-channel physiological monitors, modeling high-dimensional dependencies has become a persistent challenge. Traditional methods often assume variable independence or linearity, limiting their scalability and sensitivity to complex inter-series correlations. Deep learning architectures have introduced multivariate fusion and attention mechanisms, yet these methods still struggle with feature explosion and computational inefficiency when scaling to hundreds or thousands of correlated streams. Additionally, identifying which dimensions contribute meaningfully to anomalies remains difficult, especially when correlations vary dynamically over time. Future research must develop dimensionally adaptive and interpretable models that integrate feature selection, temporal alignment, and correlation modeling into a unified framework capable of operating efficiently in large-scale environments.

7.3 Generalizability and Transferability

Despite major advances in algorithmic sophistication, most anomaly detection models remain domain-dependent and fail to generalize across diverse datasets or applications. Models trained on financial sequences, for instance, rarely transfer effectively to medical or industrial data due to differences in sampling rates, temporal patterns, and noise characteristics. Techniques such as domain adaptation, meta-learning, and zero- or few-shot learning show promise in bridging this gap but are still in nascent stages for time-series anomaly detection. The development of generalizable feature spaces and self-supervised pretraining methods could provide universal temporal embeddings transferable across domains. Achieving true generalization requires designing algorithms that learn domain-invariant representations and dynamically adjust to new contexts without extensive retraining, thus enhancing adaptability and reducing annotation costs.

7.4 Concept Drift and Dynamic Environments

Real-world time-series data are rarely stationary; their statistical properties evolve continuously due to environmental changes, system aging, or human intervention. This concept drift poses a critical challenge because models trained on historical data may rapidly degrade as distributions shift. Static models often underperform in streaming contexts where anomalies emerge under new or unseen conditions. Addressing this issue demands adaptive learning frameworks capable of incremental updates and online retraining without catastrophic forgetting. Emerging solutions include ensemble-based detectors that reweight models over time and hybrid methods combining online optimization

with memory replay. Developing standardized protocols for drift detection and performance evaluation in dynamic environments will be essential for building resilient and continuously learning anomaly detection systems.

7.5 Reproducibility and Evaluation Inconsistency

The reproducibility crisis, long observed in broader AI research, has also permeated time-series anomaly detection. Variability in preprocessing pipelines, inconsistent metric definitions, and selective result reporting impede fair comparison across studies. Metrics such as F1, ROC-AUC, and Precision-Recall AUC are often used inconsistently, while range-based and latency-aware metrics remain underexplored despite their suitability for temporal evaluation. Furthermore, the scarcity of open-source implementations and standardized experimental protocols undermines transparency and slows cumulative progress. Addressing these limitations requires establishing unified reporting standards, adopting reproducible experimental frameworks, and mandating code and dataset sharing for publication. Such measures will enhance credibility, facilitate cross-domain benchmarking, and accelerate the translation of research into reliable real-world systems.

8. Toward a Unified Taxonomy and Future Directions

To consolidate the growing diversity of methodologies in time-series anomaly detection, this review proposes a unified, multidimensional taxonomy essential for organizing research and guiding model selection. Unlike previous surveys that classify techniques along a single analytical axis, the proposed taxonomy integrates four complementary dimensions to provide a comprehensive and comparative framework. The first axis represents supervision levels, progressing from unsupervised approaches that rely on reconstruction or prediction errors, to semi-supervised frameworks using limited labeled anomalies, and fully supervised models trained on extensive annotated datasets. The second axis classifies methods by their analytical paradigm: forecasting-based models that detect deviations between predicted and observed values, reconstruction-based architectures that identify irregularities through encoder–decoder mismatches, density-based models estimating data likelihood, and hybrid approaches integrating multiple strategies for improved robustness. The third axis differentiates anomaly types, including point anomalies reflecting isolated deviations, contextual anomalies that depend on temporal or environmental context, and collective anomalies manifesting as abnormal subsequences. The final axis pertains to deployment settings, ranging from batch analysis of historical data to real-time streaming and resource-constrained edge deployment. Integrating these four axes yields the proposed unified taxonomy, which captures both algorithmic structure and practical constraints, offering a consistent framework for evaluating, benchmarking, and transferring anomaly detection methods across diverse domains. This taxonomy establishes a conceptual foundation for developing transparent, efficient, and domain-agnostic frameworks in time-series anomaly detection. The proposed unified taxonomy categorizes time-series anomaly detection methods across four key analytical dimensions, enabling systematic comparison and model selection (Table 4, Figure 2).

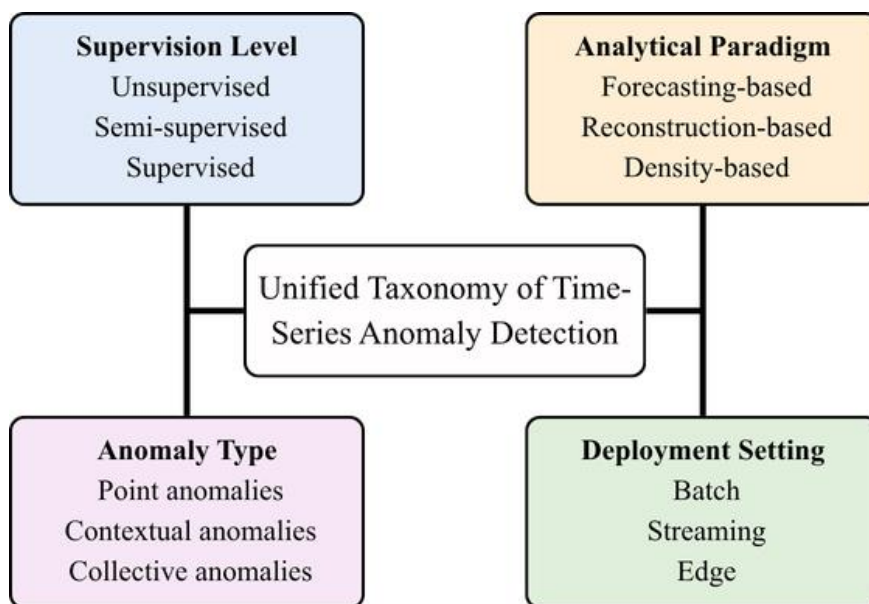


Figure 2: Unified taxonomy of time-series anomaly detection across four analytical dimensions.

Axis	Categories	Examples
Supervision	Unsupervised, Semi-supervised, Supervised	AE, LSTM-AE, OC-SVM, Labeled RNN
Paradigm	Forecasting, Reconstruction, Density-based, Hybrid	ARIMA, LSTM, VAE, TadGAN, ARIMA-LSTM
Anomaly Type	Point, Contextual, Collective	Outlier, Seasonal deviation, Pattern drift
Deployment	Batch, Streaming, Edge/Federated	Offline ARIMA, Online TCN, FedSW-TSAD

The future of time-series anomaly detection lies in the development of interpretable, scalable, and domain-agnostic models capable of adapting to heterogeneous data sources without extensive retraining. Achieving this will require unifying explainable architectures with efficient learning paradigms to ensure transparency and trust in decision-making. Establishing standardized benchmarks and open evaluation protocols will enhance reproducibility and allow fair comparisons across models and datasets. The integration of foundation models pre-trained on massive temporal corpora holds potential to extend generalization capabilities across domains while minimizing data scarcity constraints. Finally, incorporating human-centered feedback loops, where domain experts refine, validate, and contextualize algorithmic outputs, will transform anomaly detection into an interactive and continuously improving analytical ecosystem that aligns technical precision with real-world interpretability and ethical accountability. These guiding principles, transparency, efficiency, and domain-agnostic design, collectively define the foundation for next-generation anomaly detection frameworks (Figure 3).

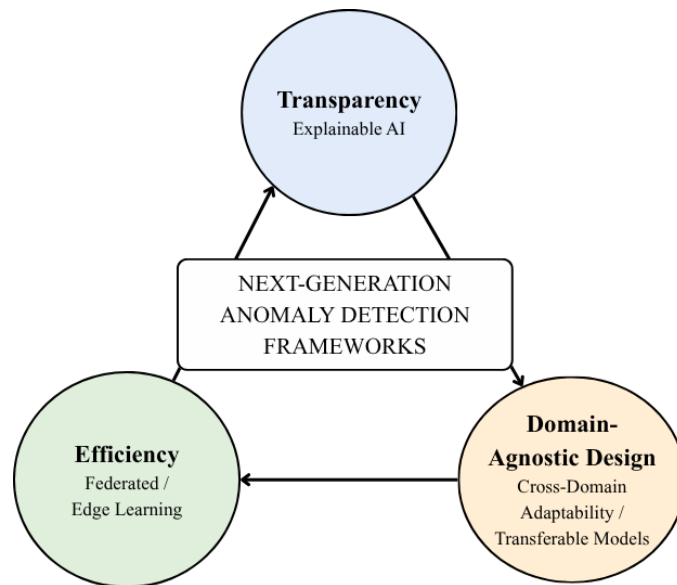


Figure 3: Conceptual framework connecting transparency, efficiency, and domain-agnostic design in next-generation anomaly detection systems

9. Conclusion

Time-series anomaly detection has progressed from interpretable statistical baselines to flexible machine learning and advanced deep architectures that learn temporal structure directly from data. Classical forecasting and decomposition methods provided clarity and rigor, while traditional learning approaches expanded analytical capacity to nonlinear and multivariate settings through clustering, isolation, and margin-based decision boundaries. Recent neural models have unified sequence forecasting, reconstruction, and attention mechanisms to capture long-range and cross-variable dependencies, enabling robust detection across increasingly complex data regimes. The field has also evolved beyond algorithms toward deployable systems that integrate benchmarking, evaluation, and operational feedback. Sustained progress now depends on transparency, adaptability, and domain-agnostic design. Transparency advances through explainable attention, saliency, and attribution techniques that render detection decisions auditable in safety-critical environments. Adaptability emerges from online and federated training paradigms that preserve privacy while addressing data drift and heterogeneity at the edge. Domain-agnostic design enhances generalization through standardized, drift-aware benchmarks, consistent reporting protocols, and reusable foundation models that transfer knowledge across domains with minimal annotation. The unified taxonomy presented in this review provides a coherent structure for mapping these methodological developments and identifying future research priorities. It supports the creation of transparent, efficient, and domain-agnostic anomaly detection frameworks capable of adapting across diverse data sources and operational settings. The next generation of anomaly detection systems will be defined by the convergence of deep learning with explainable interfaces and privacy-aware federated training. This convergence will yield models that are accurate, interpretable, and resilient, delivering timely and trustworthy insights across finance, healthcare, manufacturing, cybersecurity, and environmental monitoring. It will also foster an open evaluation culture that strengthens comparability and reproducibility, ensuring that domain-agnostic frameworks remain robust under evolving data and deployment constraints.

References

1. Alabdulatif, A. (2025). GuardianAI: Privacy-preserving federated anomaly detection with differential privacy. *Array*, 26, 100381.

2. Alamr, A., & Artoli, A. (2023). Unsupervised transformer-based anomaly detection in ECG signals. *Algorithms*, 16(3), 152.
3. Alghieth, M. (2025). DeepECG-Net: a hybrid transformer-based deep learning model for real-time ECG anomaly detection. *Scientific Reports*, 15(1), 20714.
4. Ali, T. (2024). *Next-generation intrusion detection systems with LLMs: real-time anomaly detection, explainable AI, and adaptive data generation* (Master's thesis, T. Ali).
5. Braei, M., & Wagner, S. (2020). Anomaly detection in univariate time-series: A survey on the state-of-the-art. *arXiv preprint arXiv:2004.00433*.
6. Chen, L., Xu, Y., Li, M., Hu, B., Guo, H., & Liu, Z. (2025). Privacy-preserving lightweight time-series anomaly detection for resource-limited Industrial IoT edge devices. *IEEE Transactions on Industrial Informatics*.
7. Choi, K., Yi, J., Park, C., & Yoon, S. (2021). Deep learning for anomaly detection in time-series data: Review, analysis, and guidelines. *IEEE access*, 9, 120043-120065.
8. Desani, N. R., & Chittibala, D. R. (2021). Adaptive Machine Learning Models for Real-Time Anomaly Detection in Streaming Data. *Int. J. Inf. Technol. Manag. Inf. Syst*, 12, 57-62.
9. Geiger, A., Liu, D., Alnegheimish, S., Cuesta-Infante, A., & Veeramachaneni, K. (2020, December). Tadgan: Time series anomaly detection using generative adversarial networks. In *2020 IEEE international conference on big data (big data)* (pp. 33-43). IEEE.
10. Gugulothu, N., Malhotra, P., Vig, L., & Shroff, G. (2018, July). Sparse neural networks for anomaly detection in high-dimensional time series. In *AI4IOT workshop in conjunction with ICML, IJCAI and ECAI* (pp. 1551-3203).
11. Guo, Y., Ji, T., Wang, Q., Yu, L., Min, G., & Li, P. (2020). Unsupervised anomaly detection in IoT systems for smart cities. *IEEE Transactions on Network Science and Engineering*, 7(4), 2231-2242.
12. Gupta, D., Kayode, O., Bhatt, S., Gupta, M., & Tosun, A. S. (2021, December). Hierarchical federated learning based anomaly detection using digital twins for smart healthcare. In *2021 IEEE 7th international conference on collaboration and internet computing (CIC)* (pp. 16-25). IEEE.
13. Hamon, M., Lemaire, V., Nair-Benrekia, N. E. Y., Berlemont, S., & Cumin, J. (2025). Unsupervised Feature Construction for Anomaly Detection in Time Series--An Evaluation. *arXiv preprint arXiv:2501.07999*.
14. Hao, W., Yang, T., & Yang, Q. (2021). Hybrid statistical-machine learning for real-time anomaly detection in industrial cyber-physical systems. *IEEE Transactions on Automation Science and Engineering*, 20(1), 32-46.
15. Hoeltgebaum, H., Adams, N., & Fernandes, C. (2021). Estimation, forecasting, and anomaly detection for nonstationary streams using adaptive estimation. *IEEE Transactions on Cybernetics*, 52(8), 7956-7967.
16. Huang, H., Wang, P., Pei, J., Wang, J., Alexanian, S., & Niyato, D. (2025). Deep learning advancements in anomaly detection: A comprehensive survey. *IEEE Internet of Things Journal*.
17. Hundman, K., Constantinou, V., Laporte, C., Colwell, I., & Soderstrom, T. (2018, July). Detecting spacecraft anomalies using lstms and nonparametric dynamic thresholding. In *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining* (pp. 387-395).
18. Inuwa, M. M., & Das, R. (2024). A comparative analysis of various machine learning methods for anomaly detection in cyber attacks on IoT networks. *Internet of Things*, 26, 101162.
19. Iqbal, A., & Amin, R. (2024). Time series forecasting and anomaly detection using deep learning. *Computers & Chemical Engineering*, 182, 108560.
20. Iqbal, A., & Amin, R. (2025). An efficient mechanism for time series forecasting and anomaly detection using explainable artificial intelligence. *The Journal of Supercomputing*, 81(4), 523.

21. Jin, M., Koh, H. Y., Wen, Q., Zambon, D., Alippi, C., Webb, G. I., ... & Pan, S. (2024). A survey on graph neural networks for time series: Forecasting, classification, imputation, and anomaly detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.
22. Karadayi, Y., Aydin, M. N., & Öğrenci, A. S. (2020). Unsupervised anomaly detection in multivariate spatio-temporal data using deep learning: early detection of COVID-19 outbreak in Italy. *Ieee Access*, 8, 164155-164177.
23. Kim, J., Kang, H., & Kang, P. (2023). Time-series anomaly detection with stacked Transformer representations and 1D convolutional network. *Engineering Applications of Artificial Intelligence*, 120, 105964.
24. Kumar, K. S., Nair, S. A. H., Roy, D. G., Rajalingam, B., & Kumar, R. S. (2021). Security and privacy-aware artificial intrusion detection system using federated machine learning. *Computers & Electrical Engineering*, 96, 107440.
25. Lee, T. J., Gottschlich, J., Tatbul, N., Metcalf, E., & Zdonik, S. (2018). Precision and recall for range-based anomaly detection. *arXiv preprint arXiv:1801.03175*.
26. Li, L., Yan, J., Wang, H., & Jin, Y. (2020). Anomaly detection of time series with smoothness-inducing sequential variational auto-encoder. *IEEE transactions on neural networks and learning systems*, 32(3), 1177-1191.
27. Li, Z., Zhu, Y., & Van Leeuwen, M. (2023). A survey on explainable anomaly detection. *ACM Transactions on Knowledge Discovery from Data*, 18(1), 1-54.
28. Malik, H., Feng, J., Shao, P., & Abduljabbar, Z. A. (2024). Improving flood forecasting using time-distributed CNN-LSTM model: a time-distributed spatiotemporal method. *Earth Science Informatics*, 17(4), 3455-3474.
29. Mandaleeka, A. V. S. S., & Bonthu, S. (2025). Hybrid Traditional and Deep Learning Approaches for Network Traffic Forecasting with Real-Time Anomaly Detection: A Comparative Study of ARIMA, LSTM, ARIM-LSTM Hybrid Model.
30. Meng, Q., Qian, H., Liu, Y., Xu, Y., Shen, Z., & Cui, L. (2023). Unsupervised representation learning for time series: A review. *arXiv preprint arXiv:2308.01578*.
31. Nizam, H., Zafar, S., Lv, Z., Wang, F., & Hu, X. (2022). Real-time deep anomaly detection framework for multivariate time-series data in industrial IoT. *IEEE Sensors Journal*, 22(23), 22836-22849.
32. Nogami, K., Tamura, H., & Tanaka, G. (2025). Federated learning with reservoir state analysis for time series anomaly detection. *arXiv preprint arXiv:2502.05679*.
33. Paparrizos, J., Boniol, P., Liu, Q., & Palpanas, T. (2025, August). Advances in time-series anomaly detection: Algorithms, benchmarks, and evaluation measures. In *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V. 2* (pp. 6151-6161).
34. Pereira, J., & Silveira, M. (2019). Unsupervised representation learning and anomaly detection in ECG sequences. *International Journal of Data Mining and Bioinformatics*, 22(4), 389-407.
35. Puder, A., Zink, M., Seidel, L., & Sax, E. (2024). Hybrid anomaly detection in time series by combining kalman filters and machine learning models. *Sensors*, 24(9), 2895.
36. Rasul, I., Shaboj, S. I., Rafi, M. A., Miah, M. K., Islam, M. R., & Ahmed, A. (2024). Detecting Financial Fraud in Real-Time Transactions Using Graph Neural Networks and Anomaly Detection. *Journal of Economics, Finance and Accounting Studies*, 6(1), 131-142.
37. Rojat, T., Puget, R., Filliat, D., Del Ser, J., Gelin, R., & Díaz-Rodríguez, N. (2021). Explainable artificial intelligence (xai) on timeseries data: A survey. *arXiv preprint arXiv:2104.00950*.
38. Ruff, L., Vandermeulen, R., Goernitz, N., Deecke, L., Siddiqui, S. A., Binder, A., ... & Kloft, M. (2018, July). Deep one-class classification. In *International conference on machine learning* (pp. 4393-4402). PMLR.
39. Schmidl, S., Wenig, P., & Papenbrock, T. (2022). Anomaly detection in time series: a comprehensive evaluation. *Proceedings of the VLDB Endowment*, 15(9), 1779-1797.

40. Sharif, M. H., Gupta, K., Mohammed, M. A., & Jiwani, N. (2022). Anomaly detection in time series using deep learning. *International Journal of Engineering Applied Sciences and Technology*, 7(6), 296-305.
41. Shaukat, K., Alam, T. M., Luo, S., Shabbir, S., Hameed, I. A., Li, J., ... & Javed, U. (2021). A review of time-series anomaly detection techniques: A step to future perspectives. In *Future of Information and Communication Conference* (pp. 865-877). Springer, Cham.
42. Sørnbø, S., & Ruocco, M. (2024). Navigating the metric maze: A taxonomy of evaluation metrics for anomaly detection in time series. *Data Mining and Knowledge Discovery*, 38(3), 1027-1068.
43. Sun, G., Yin, C., Xia, T., Lu, Y., & Mao, J. (2024, November). An Improved ARIMA Based Anomaly Detection Method for Time Series Data. In *2024 IEEE 8th Conference on Energy Internet and Energy System Integration (EI2)* (pp. 5132-5138). IEEE.
44. Tibau, X. A., Reimers, C., Requena-Mesa, C., & Runge, J. (2021). Spatio-temporal Autoencoders in Weather and Climate Research. *Deep Learning for the Earth Sciences: A Comprehensive Approach to Remote Sensing, Climate Science, and Geosciences*, 186-203.
45. Truong, C., Oudre, L., & Vayatis, N. (2020). Selective review of offline change point detection methods. *Signal Processing*, 167, 107299.
46. Wang, S., Tao, J., Jiang, Q., Chen, W., Qin, C., & Liu, C. (2024). A digital twin framework for anomaly detection in industrial robot system based on multiple physics-informed hybrid convolutional autoencoder. *Journal of Manufacturing Systems*, 77, 798-809.
47. Xu, H., Pang, G., Wang, Y., & Wang, Y. (2023). Deep isolation forest for anomaly detection. *IEEE Transactions on Knowledge and Data Engineering*, 35(12), 12591-12604.
48. Xu, R., Miao, H., Wang, S., Yu, P. S., & Wang, J. (2024, August). PeFAD: a parameter-efficient federated framework for time series anomaly detection. In *Proceedings of the 30th ACM SIGKDD conference on knowledge discovery and data mining* (pp. 3621-3632).
49. Yang, S., Yang, D., Qu, J., Fang, Z., Hu, X., Xu, Q., & Liu, J. Gtf-Net: A Dynamic Graph–Temporal Fusion Intrusion Detection Model Tailored for Can Networks. *Available at SSRN 5392983*.
50. Yang, X., Qi, X., & Zhou, X. (2023). Deep learning technologies for time series anomaly detection in healthcare: A review. *Ieee Access*, 11, 117788-117799.
51. Zamanzadeh Darban, Z., Webb, G. I., Pan, S., Aggarwal, C., & Salehi, M. (2024). Deep learning for time series anomaly detection: A survey. *ACM Computing Surveys*, 57(1), 1-42.
52. Zhang, K., Jiang, Y., Seversky, L., Xu, C., Liu, D., & Song, H. (2021, October). Federated variational learning for anomaly detection in multivariate time series. In *2021 IEEE International Performance, Computing, and Communications Conference (IPCCC)* (pp. 1-9). IEEE.
53. Zhang, X., Zhao, H., Zhang, W., Cao, S., Sun, H., & Zhang, B. (2025). FedSW-TSAD: SWGAN-Based Federated Time Series Anomaly Detection. *Sensors*, 25(13), 4014.
54. Zhu, F., Ning, D. J., Wang, Y., & Liu, S. (2021, December). A novel cost-sensitive capsule network for audit fraud detection. In *2021 20th International Conference on Ubiquitous Computing and Communications (IUCC/CIT/DSCI/SmartCNS)* (pp. 549-556). IEEE.