# Automated Neural Network Auditing (ANNA) for Energy Market Settlements: A Machine and Reinforcement Learning-Based Approach

September 23, 2025


## Abstract

Abstract – Energy market settlements in ISOs such as CAISO and MISO have grown increasingly complex with the expansion of EIMs and automated transaction processes. Traditional rule-based, manual auditing is inefficient, error-prone, and unable to manage the massive data volumes these markets generate. To address this, we propose Automated Neural Network Auditing (ANNA), a hybrid machine and reinforcement learning framework for scalable anomaly detection. ANNA applies neural networks for anomaly identification and reinforcement learning for adaptive, real-time auditing, incorporating statistical methods, K-nearest Neighbors (KNN), and Deep Q-Networks (DQN). Using three years of CAISO settlement data covering 32 charge codes and over 63 million transactions, ANNA is evaluated against traditional statistical and clustering models on precision, recall, efficiency, and false negative reduction. Preliminary results show ANNA outperforms conventional approaches by uncovering hidden relationships in settlement data, enabling more accurate, compliant, and efficient audits.



**Keywords:** Energy Market Settlements, Automated Auditing, Neural Networks, Anomaly Detection, Reinforcement Learning, Machine Learning.


# I. INTRODUCTION

## A. Background and Motivation

Increased participation in ISOs, EIMs, and EDAM has made energy market settlements more complex; ISOs such as CAISO and MISO now process billions in highly granular transactions that demand verification, transparency, and compliance [1]. The sheer volume creates substantial challenges for auditing and anomaly detection [2], rendering traditional rule-based, manual verification approaches infeasible at scale [3]. Fixed rule-based and statistical frameworks lack adaptability, missing anomalies amid dynamic prices, new charge codes, shifting policies, and hidden relationships [4], [5] and their labor intensity makes them impractical for large-scale operations [6]. These inefficiencies translate into financial losses, compliance risks, inaccurate settlements, and missed cost-recovery opportunities for participants and regulators [7].

## B. Problem Statement

Traditional methods struggle to audit modern energy market settlements at scale and frequency; fixed-threshold statistical models tied to preset confidence intervals miss subtle, context-dependent anomalies [8], and their limited scalability prevents real-time processing of massive data streams [9]. Key challenges include: (1) sheer transaction volume, millions of records at varying intervals, making manual audits impractical [10]; (2) complex charge codes with variable price structures, interval pricing, and time-based adjustments that confound legacy checks [11]; (3) poor adaptability to new resources, codes, and pricing regimes [12]; and (4) high false-positive and omission rates that burden analysts and forfeit cost-recovery opportunities [13]. An AI-driven auditing system is therefore needed to detect anomalies, automate verification, and reduce human limitations.

## C. Research Objectives

The objective of this research is to develop Automated Neural Network Auditing (ANNA), an AI-powered system to improve the efficiency, accuracy, and scalability of power-market settlement audits. ANNA layers statistical benchmarking, machine learning (ML), and reinforcement learning (RL) to enable adaptive auditing, using neural networks to learn patterns from historical settlements and predict discrepancies before they affect financials [14].to

Key research objectives include:

- Real-time detection of settlement anomalies with ML.
- Strengthened compliance monitoring under ISO rules.
- Reduced financial risk via early error detection.
- Higher detection accuracy using Deep Q-Network (DQN) RL.
- Fewer false positives through K-Nearest Neighbors (KNN) and historical benchmarking.

### D. Contributions

This work: (1) introduces an AI-driven, scalable alternative to rule-based auditing; (2) integrates neural networks, clustering, and reinforcement learning with statistical benchmarking to boost anomaly-detection precision and adaptability; (3) validates the approach on three years of CAISO data spanning 32 charge codes and over 63 million WEIM settlement records [15]; (4) optimizes computation by prioritizing high-risk transactions for human review; and (5) delivers a modular architecture deployable across multiple ISOs, balancing accuracy and adaptability in diverse markets [16].

### E. Structure of the Paper

Section II condenses related work on rule-based auditing, ML anomaly detection, and RL in financial markets. Section III presents ANNA data preprocessing, essential features, and training. Section IV covers experimental setup, dataset, and evaluation metrics. Section V reports results comparing statistical, clustering, and RL methods. Section VI outlines challenges, limitations, and future improvements. Section VII concludes with key findings and implications for power-market anomaly detection. Overall, the study demonstrates a practical transition from rule-based audits to an AI-driven framework that improves accuracy, efficiency, and compliance in large-scale settlements.

## II. RELATED WORK

### A. Existing Auditing Techniques in Energy Markets

Traditional settlement auditing relies on rule checks, statistical thresholds, and manual review, but these rigid methods struggle to adapt to evolving market conditions [17]. Threshold techniques (e.g., Z-scores, confidence intervals) often fail to separate normal volatility from true discrepancies [18]. As transaction volumes and charge-code complexity have grown, driven in part by new technologies, manual verification has become impractical and risky, leading to inefficiencies, higher costs, and missed anomalies in high-volume markets [5]. Consequently, researchers have turned to automated approaches using statistical and machine learning methods.

### B. Machine Learning for Anomaly Detection

Machine learning (ML) for financial and power-market anomaly detection is increasingly prevalent: models learn patterns from historical transactions to flag outliers and improve accuracy over time [4]. Unsupervised methods like K-means, Isolation Forests, One-Class SVMs can surface deviations without labels and scale to high volumes, but struggle with high-dimensional data and shifting market regimes, leading to false positives and limited adaptability [2], [13]. Supervised approaches (neural networks, gradient boosting, ensembles) boost settlement detection accuracy but require labeled datasets that are often impractical in power markets [9]. Deep learning (e.g., Autoencoders, LSTMs) captures temporal dependencies and subtle, context-dependent anomalies, yet real-time deployment is constrained by heavy compute and large training data demands [1], [3].

### C. Reinforcement Learning in Auditing and Anomaly Detection

Reinforcement learning (RL) offers a promising, adaptive alternative for anomaly detection in financial and power-market auditing, dynamically adjusting thresholds to market behavior [12]. Deep Q-Networks (DQN), combining Q-learning with deep net optimize decisions in changing environments and have shown strong results in fraud, credit risk, and market anomaly detection [7], [13]. In settlements, RL can continually update its strategy from new transactions for real-time auditing, classifying records as valid or anomalous while using rewards to prioritize high-risk cases and curb false positives [9]. Still, applications at large ISO scale remain limited, underscoring the need for further study [1].

### D. Gaps in Existing Research and Motivation for ANNA

Despite advances in ML and RL, gaps remain for ISO settlement anomaly detection: most work targets generic financial fraud, not ISO charge codes or market structures [15], and traditional ML lacks adaptability to evolving codes, resource mixes, volatility, and regulatory change [16]. ANNA addresses this by hybridizing statistical benchmarking, unsupervised clustering, and RL to boost accuracy, cut false positives, and enable real-time auditing; RL further adapts decisions as new market data arrives [17]. To move beyond synthetic/small-scale studies, we evaluate ANNA on three years of CAISO data with over 63 million charge-code transactions at 5- and 15-minute intervals across multiple resources, assessing scalability and real-world feasibility [18]. In doing so, ANNA advances automated auditing to enhance compliance, reduce financial risk, and improve verification accuracy.

## III. PROPOSED METHODOLOGY

The Automated Neural Network Auditing (ANNA) framework is designed to provide intelligent, scalable, and adaptive anomaly detection to power market settlements. This section outlines the methodological framework of ANNA, including the data preprocessing procedures, feature engineering strategies, and the machine learning model architecture and training protocols applied to neural network development and reinforcement learning-based anomaly detection.

## A. Dataset Summary

We evaluate ANNA on real CAISO settlements from Jan 2021–Dec 2023, a period spanning resource shifts, seasonal/market conditions, and policy updates. The corpus contains over 63 million transactions across more than 32 charge codes, with metadata including Position Name, Settlement Type, Trade/Invoice Date, Hour Ending, Charge/Determinant, Amount/Value, sufficient for learning contextual behavior and isolating anomalies.

Representative charge codes (each with distinct temporal/volumetric patterns):

- SettlementIntervalRealTimeLmp (5-min LMP),
- SettlementIntervalRealTimeUie (unaccounted interval energy),
- DailyInitialMarket, DailyRecalcMarket (daily forecast/recalc),
- DailyRerunMarket, DailyRerunMarket512 (policy/pricing reruns).

## B. Data Preprocessing

Charge codes exhibit distinct temporal and volumetric patterns, enabling the model to learn baseline settlements and corrective adjustments. To ensure quality and model compatibility, we: (1) filtered out "Shadow" entries to focus on finalized ISO data; (2) converted non-numeric Amount values to floats and normalized numerics; (3) dropped invalid/missing records; (4) encoded categoricals; and (5) derived time features (month, weekday, hour) from timestamps.

For localized anomaly detection, data were partitioned by Settlement Type and Position Name, then grouped by these contexts so the model could capture participant and mechanism-specific behavior, improving interpretability and anomaly resolution. The resulting, large-scale CAISO dataset (millions of rows over three years) provides sufficient volume, breadth, and granularity to evaluate ANNA's performance and generalizability across market conditions, participants, resources, and charge structures.

## C. Feature Engineering

The feature engineering employed to transform the raw settlement records into structured, learnable representations include:

- Temporal indices: Derived from parsed Trade Date and Invoice Date fields, including Month, Day of Week, and Hour Ending.
- Categorical encodings: Like Charge / Determinant Name and Settlement Type, encoded using one-hot or label encodings depending on model requirements.
- Benchmark statistics: Local rolling mean and standard deviation were calculated for each charge code within a group to normalize transactional behavior over time.

These essential features, enabling neural network and reinforcement learning agents to distinguish between normal operations and anomalies in settlement patterns, were created through the feature engineering employed.

## D. Neural Network Architecture

ANNA uses a hybrid pipeline statistical benchmarking, unsupervised clustering, reinforcement learning, with a Deep Q-Network (DQN) as the primary agent. In a Gymnasium-style environment, each transaction is a state and the action is {normal, anomalous}. The reward reflects alignment with deviation signals (e.g., departures from learned confidence bounds). DQN learning follows the Bellman update, optimizing expected return over actions to identify anomalous settlements efficiently.

The decision-making process is improved through iterative updates of the Q-table using the Bellman equation:

$$Q(s,a) \leftarrow Q(s,a) + \alpha \left[ r + \gamma \max_{a'} Q(s',a') - Q(s,a) \right]$$

Where $\alpha$ is the learning rate, $\gamma$ is the discount factor, $r$ is the reward, and $s$ and $a$ are the respective state and action. Anomalies are identified when the model encounters high temporal-difference errors, indicating a transaction does not conform to the expected behavioral policy.

To replace the static lookup table, the DQN trains a neural network to approximate Q-function. The network takes the current state as input and outputs Q-values for each possible action. The loss function minimized during training is based on the squared difference of the predicted Q-value and the target computed from the Bellman equation:

$$\text{Loss} = \left[ Q(s,a) - \left( r + \gamma \max_{a'} Q(s',a') \right) \right]^2$$

Experienced replay stabilizes learning by storing previous state-action-reward sequences to be sampled randomly to train the network. A separate target network is periodically updated to reduce instability during training.

With the incorporation of DQN into the framework, ANNA can detect subtle and complex evolving

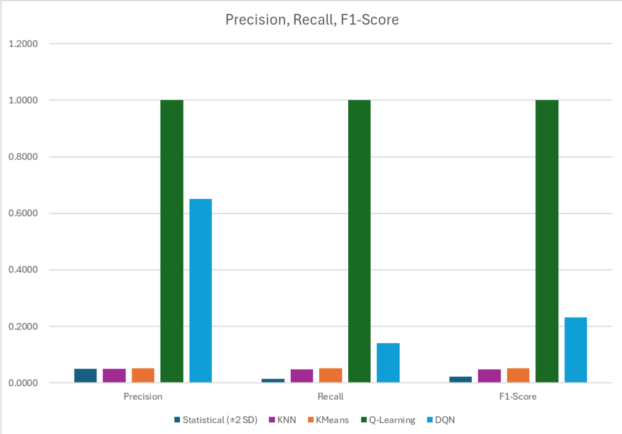


Figure 1: Precision, Recall, F1-Score.

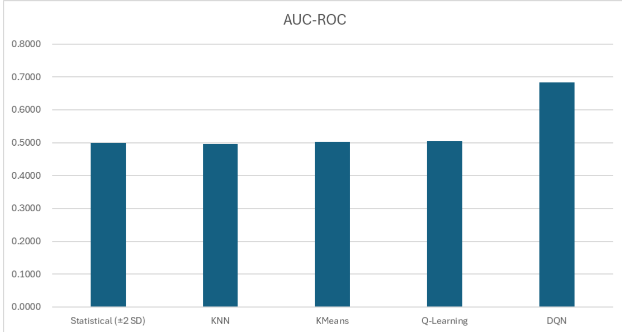


Figure 2: Model Accuracy.

anomaly patterns, adapt to new data distributions, and improve its anomaly detection policy with each learning episode.

## E. Training Procedure

Models were trained per Position Name × Settlement Type to enable localized learning, with multiple RL episodes and epsilon-greedy exploration decay so the policy adapts to evolving transaction patterns; performance feedback relied on precision, recall, and delta-based metrics by comparing model flags to statistical outliers, and forecasted amounts were retained to support post-audit reconciliation; the system was implemented in Python using NumPy, Pandas, scikit-learn, PyGAM, and Gymnasium, with a modular architecture designed for scalability and portability across energy markets and audit environment.

# IV. Experiment Results

Evaluation benchmarks covered predictive accuracy and computational efficiency: detection fidelity via Precision, Recall, F1, Accuracy; discrimination across thresholds via AUC-ROC; and system-level metrics (execution time, memory) to assess scalability and deployability under realistic, high-volume, time-sensitive settlement conditions. This multi-metric analysis grounds conclusions in both predictive power and operational feasibility.

## A. Detection Precision, Recall, and F1-Score

Fig. 1: Precision, Recall, F1-Score.

Fig. 1. compares precision, recall, and F1 for five methods: Statistical (±2), KNN, KMeans, Q-Learning, and DQN. Q-Learning reports perfect scores (Precision/Recall/F1 = 1.000), indicating likely overfitting with poor robustness to unseen data. DQN provides the most credible profile: Precision 0.6515, Recall 0.1415, F1 0.2326, balancing sensitivity and specificity for real-time auditing. Statistical (±2) performs poorly (P 0.0500 / R 0.0146 / F1 0.0226), underscoring rigidity in high-variance settings. KNN/KMeans are similarly weak (F1 0.0489/0.0516), reflecting limits in high-dimensional, noisy data. Overall, DQN is the most effective practical method; the others either overfit (Q-Learning) or lack adaptability (clustering, ±2), reinforcing the need for adaptive, learning-based systems.

## B. Accuracy Comparison Across Models

Fig. 2: Model Accuracy.

The comparative accuracy achieved by the five anomaly detection models evaluated within the ANNA framework are presented in Fig 2. Accuracy comparisons can mislead under severe class imbalance, as high scores may reflect majority-class bias rather than true detection quality. Q-Learning posts 1.000 accuracy (and perfect P/R/F1), signaling overfitting and poor generalization, warranting k-fold and OOD tests. DQN is next at 0.9565 and, unlike Q-Learning, pairs this with credible Precision 0.6515 / Recall 0.1415, indicating meaningful anomaly capture without excessive false positives. Statistical (±2) reaches 0.9368 but remains rigid with weak detection fidelity, while KMeans/KNN trail at 0.9045/0.9042, consistent with misclassification in high-dimensional, noisy data lacking temporal/context features. Overall, DQN provides the most operationally meaningful accuracy; Q-Learning's perfection reflects memorization, and clustering/statistical baselines underperform in dynamic, imbalanced settlements.

## C. Receiver Operating Characteristic (AUC-ROC) Evaluation

Fig. 3: Area Under the ROC.

Fig 3. AUC-ROC summarizes discrimination across thresholds (1.0 = strong separation; 0.5 random). DQN leads with AUC 0.6829, indicating effective ranking for audit triage. KMeans/KNN hover near random

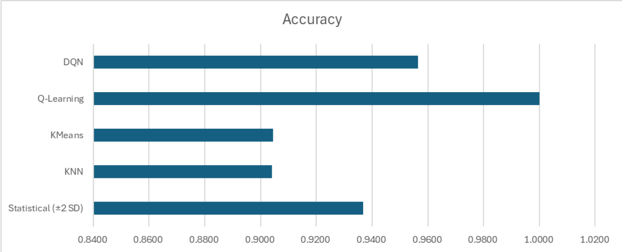


Figure 3: Area Under the ROC.

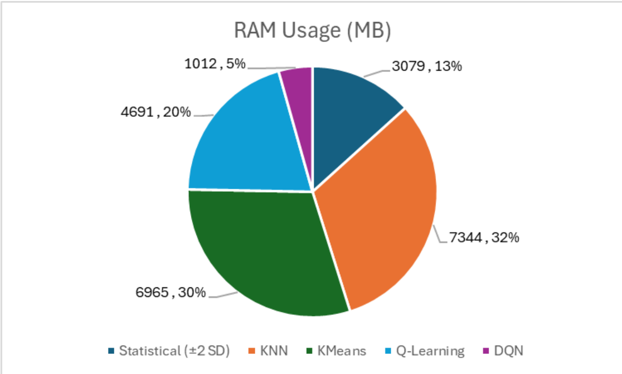


Figure 4: RAM Usage (MB).

(0.5034 / 0.4966), revealing limited boundary flexibility in high-dimensional data. Q-Learning, despite perfect accuracy/precision/recall/F1, scores only 0.5036 AUC, reinforcing overfitting and weak rank ordering. Statistical (±2) is 0.5000, offering no meaningful ranking. Overall, DQN provides the needed threshold flexibility and ranking granularity for risk-sensitive auditing.

### D. Memory Utilization and Computational Efficiency

Fig. 4: RAM Usage (MB).

Memory is critical for real-time, high-throughput settlement auditing. Fig. 4. presents the comparative RAM usage where DQN is most efficient at 1,011.9 MB, pairing a small footprint with strong detection ideal for containerized or quota-bound deployments. KNN/KMeans are heaviest (7,343.7 MB / 6,965.4 MB) due to pairwise distances, neighbor indices, and iterative convergence, scaling poorly with data volume and dimensionality. Statistical (±2) and Q-Learning are mid-range (3,079.1 MB / 4,691.1 MB), driven by rolling caches and tabular Q-value growth, respectively. Overall, DQN offers the best resource profile for distributed/edge auditing.

### E. Runtime Performance

Fig. 5: Results Summary.

Runtime is critical for real-time settlement auditing. Fig. 5. depicts the execution times where the DQN is fastest at 0.0312s (vectorized ops, batching, efficient policy inference). The Statistical (±2) baseline is slowest at 73.51s due to rolling SD/confidence calculations on dense data. KMeans/KNN take 1.1865s / 0.9994s, reflecting repeated distance computations and poor iterative scaling, fine for batch audits, but not streaming. Q-Learning is also fast but less generalizable, whereas DQN offers the best balance of low latency and detection performance for high-frequency, automated auditing.

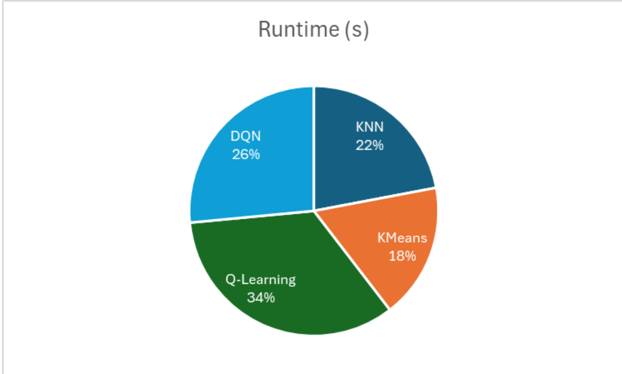


Figure 5: Results Summary.

## V. Discussion and Future Work

### A. Key Takeaways

Across five methods (Statistical ±2, KNN, KMeans, Q-Learning, DQN), ANNA shows clear advantages, with DQN delivering the best balance of detection, runtime, and resource use. The Statistical (±2) baseline is weak (P 0.0500 / R 0.0146 / F1 0.0226) and slow (73.51s). KNN/KMeans offer only slight gains (F1 0.0489/0.0516) but are memory-heavy (7.34 GB / 6.97 GB), limiting real-time use. Q-Learning achieves perfect metrics (P/R/F1 = 1.000) yet likely overfits; it's fast (0.040s) with moderate memory (4.69 GB), suiting controlled/static settings. DQN emerges as most operational: Precision 0.651, Recall 0.1415, F1 0.2326, AUC 0.6829, with the lowest RAM (1.01 GB) and 0.0312s runtime, ready for real-time, high-throughput auditing.

### B. Internal and External Threats to Study Results

Despite strong results, ANNA faces practical constraints that affect scalability, reliability, and adoption: (1) Compute: training DQN on multi-year, high-resolution streams is resource-intensive; thousands of localized models may require parallelism, model pruning, or distributed training to keep runtimes manageable. (2) Data privacy: effective anomaly detection depends on granular, accurately labeled settlement data that may include sensitive PII; approaches like federated learning and differential privacy can mitigate sharing and governance risks. (3) Regulation interpretability: auditors and regulators need transparent, explainable rationales; even with stored forecasts and deviation scores, the system can appear "black box," so standardized AI audit trails and explainability tooling are essential. Addressing these constraints is key to making ANNA robust, defensible, and interoperable in evolving regulatory environments.

### C. Future Enhancements

ANNA's maturation from proof-of-concept to operational prototypes requires enhancement to increase accuracy, adaptability, and relevance in a large range of market contexts, including:

- CFT features. Add Complex Fourier Transforms to capture frequency-domain patterns (periodic/cyclic anomalies) that time-domain models miss; magnitude/phase features enrich clustering/RL inputs and reduce false negatives in dynamic pricing [1], [3].
- Multi-ISO expansion. Parameterize ISO-specific charge codes, intervals, and audit triggers (PJM, NYISO, ERCOT, MISO) while retaining a common detection core for scalability and interoperability.
- Blockchain provenance. Use DLT for immutable, time-stamped logs of anomalies/decisions/actions to strengthen traceability, regulatory trust, and dispute resolution.

Together, these upgrades improve ANNA's accuracy, adaptability, geographic scope, and auditability, advancing it toward a production-grade, compliant settlement-auditing framework.

## VI. Conclusions

With the energy sector's continued digital transformation, increased complexity and volume of financial settlements necessitate a transition from traditional auditing methods to more advanced tools capable of intelligently identifying anomalous behavior in real time. This paper outlines the design, implementation, and evaluation of a hybrid AI-based Automated Neural Network Auditing (ANNA) auditing system, integrating statistical benchmarking, unsupervised clustering, and reinforcement learning for transaction settlement anomaly detection. Through modular architecture and data-driven methodology, the ANNA framework addresses critical needs for scalability, accuracy, and interpretable auditing solutions in dynamic energy markets.

### A. Summary of Contributions:

- The application of Deep Q-Networks (DQNs) in a simulated Gymnasium-based environment enabled the detection of anomalous settlement transactions through reward-optimized policy learning.
- Application of statistical deviation metrics and unsupervised clustering to benchmarks and characterize traditional transactional outliers across multiple Positions and Settlement Types to improve system sensitivity to diverse patterns of anomalous behaviors.
- Reinforcement learning dynamic reward shaping to optimize anomaly detection for statistical divergence and operation relevance in support of actionable audit outcomes.
- Utilization of 2.5 million hourly settlement records across a 24-month period to train and validate the system with high-impact charge codes indicative of market volatility and proper audit priority.

The ANNA framework demonstrated strong potential in detecting complex temporal deviations traditional rule-based systems fail to capture.

### B. Real-World Impact and Industry Relevance:

The ANNA framework's real-world practical application enhances market settlements by automating the identification of settlement anomalies caused by pricing discrepancies, billing inconsistencies, and potential data manipulation. The ability to adapt to evolving market conditions and transactions is a significant improvement over the existing static compliance framework. Integrating anomaly explanation, forecasted values, and post-detection audit support enhances critical system transparency and traceability for real-world deployment.

### C. Adoption by ISOs and Regulatory Potential:

With its modular market-agnostic design, the ANNA framework has significant promise for deployment across various Independent System Operators (ISOs) such as CAISO, PJM, ERCO, NYIS, and MISO. With increased emphasis on automation, transparency, and data integrity in market oversight, ANNA is well-positioned as a foundation tool in regulatory compliance, dispute resolution, and post-settlement reconciliation. Future enhancements like blockchain integration and spectral feature analysis will solidify ANNA's role in next-generation energy auditing ecologies.

### D. Conclusion:

ANNA is a transformative step toward explainable, intelligent, operationally effective auditing in high-volume, high-stakes energy markets. ANNA offers a blueprint supporting accelerated AI adoption throughout the industry, enhancing auditing precision, reducing financial risk, and reinforcing settlement integrity.

# References


[1] Ghanim, J., & Awad, M., "An unsupervised anomaly detection in electricity consumption using reinforcement learning and time series forest-based framework," arXiv preprint arXiv:2501.00107, 2024. [Online]. Available: https://arxiv.org/abs/2501.00107

[2] Himeur, Y., Ghanem, K., & Bensaali, F., "Artificial intelligence based anomaly detection of energy consumption data: A review, current trends and new perspectives," Applied Energy, vol. 287, p. 116601, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0306261921001409

[3] Himeur, Y., Ghanem, K., & Bensaali, F., "Explaining deep learning-based anomaly detection in energy consumption: A comprehensive survey," Energy and Buildings, vol. 278, p. 112519, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378778824012933

[4] Wang, Y., Zhang, N., & Li, Y., "Reinforcement learning in deregulated energy market: A comprehensive review," Applied Energy, vol. 312, p. 118662, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0306261922014696

[5] Zhang, Y., & Wang, J., "Energy grid management system with anomaly detection and Q-learning decision modules," Computers & Electrical Engineering, vol. 103, p. 108395, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0045790623000642

[6] Telila, Y. K., Senevirathne, D., Tissera, D., Narayan, A., Capretz, M. A. M., & Grolinger, K., "Federated learning for anomaly detection in energy consumption data: Assessing the vulnerability to adversarial attacks," arXiv preprint arXiv:2502.05041, 2025. [Online]. Available: https://arxiv.org/abs/2502.05041

[7] Himeur, Y., Ghanem, K., Bensaali, F., & Amira, A., "AI-empowered methods for smart energy consumption: A review," International Journal of Precision Engineering and Manufacturing-Green Technology, vol. 10, no. 1, pp. 1–22, 2023. [Online]. Available: https://link.springer.com/article/10.1007/s40684-023-00537-0

[8] Kontokosta, C., Bonczak, B., & Duer-Balkind, M., "DataIQ – A machine learning approach to anomaly detection for energy performance data quality and reliability," Proceedings of the ACEEE Summer Study on Energy Efficiency in Buildings, 2016. [Online]. Available: https://www.aceee.org/files/proceedings/2016/data/papers/12_1139.pdf

[9] Li, H., & Zhang, X., "Abnormal detection and early warning mechanism of power market based on improved isolation forest algorithm," International Journal of Low-Carbon Technologies, 2024. [Online]. Available: https://academic.oup.com/ijlct/article/doi/10.1093/ijlct/ctae147/7745536

[10] Gallo, A., "Analysis of machine learning methods for anomaly detection of power consumption in buildings," Politecnico di Milano, 2021. [Online]. Available: https://www.politesi.polimi.it/bitstream/10589/183344/3/Final%20Version%201.pdf

[11] Wageningen University & Research, "AI-powered anomaly detection in energy markets (MSc)," n.d.. [Online]. Available: https://www.wur.nl/en/article/ai-powered-anomaly-detection-in-energy-markets-msc.htm

[12] Pan, H., Yin, Z., & Jiang, X., "High-dimensional energy consumption anomaly detection: A deep learning-based method for detecting anomalies," Energies, vol. 15, no. 17, p. 6139, 2022. [Online]. Available: https://www.mdpi.com/1996-1073/15/17/6139

[13] Kouziokas, G. N., & Perakis, K., "Enhancing resilience in complex energy systems through real-time anomaly detection: A systematic literature review," Energy Informatics, vol. 7, no. 1, pp. 1–22, 2024. [Online]. Available: https://energyinformatics.springeropen.com/articles/10.1186/s42162-024-00401-8

[14] Zhai, S., Cheng, Y., Lu, W., & Zhang, Z., "Deep structured energy based models for anomaly detection," arXiv preprint arXiv:1605.07717, 2016. [Online]. Available: https://arxiv.org/abs/1605.07717

[15] Zhang, J. E., Wu, D., & Boulet, B., "Time series anomaly detection via reinforcement learning-based model selection," arXiv preprint arXiv:2205.09884, 2022. [Online]. Available: https://arxiv.org/abs/2205.09884

[16] Wu, T., & Ortiz, J., "RLAD: Time series anomaly detection through reinforcement learning and active learning," arXiv preprint arXiv:2104.00543,

2021. [Online]. Available: https://arxiv.org/abs/2104.00543

[17] Himeur, Y., Ghanem, K., Alsalemi, A., Bensaali, F., & Amira, A., "Artificial intelligence based anomaly detection of energy consumption in buildings: A review, current trends and new perspectives," arXiv preprint arXiv:2010.04560, 2020. [Online]. Available: https://arxiv.org/abs/2010.04560

[18] Mathonsi, T., & van Zyl, T. L., "Statistics and deep learning-based hybrid model for interpretable anomaly detection," arXiv preprint arXiv:2202.12720, 2022. [Online]. Available: https://arxiv.org/abs/2202.12720