Asia Pacific Science Press

Journal of Advances in Engineering and Technology
ISSN(O):3005-7507
ISSN(P):3005-7493
Vol. 2 No.3 (2025)

Anomaly Detection in Cold Storage Systems: A Machine Learning Approach for Fault Diagnosis

David Kumar, Fatima Rahman*

School of Information Technology, University of Toronto, Canada

Corresponding author: Fatima Rahman, fatima.rahman6@utoronto.ca

Abstract: Cold storage systems play a crucial role in preserving temperature-sensitive goods. However, they are susceptible to various faults that can compromise operational efficiency and product safety. Traditional rule-based fault detection methods are limited by their rigidity and lack of adaptability. In contrast, this study introduces a machine learning (ML)-based framework for anomaly detection and fault diagnosis in cold storage environments. The proposed framework combines autoencoders for unsupervised anomaly detection with gradient boosting classifiers for supervised fault categorization. It addresses key challenges such as data imbalance, temporal drift, and sensor noise. Experimental results on an industrial cold storage dataset show that the framework achieves high fault detection accuracy, reduced false alarm rates, and strong generalization to unseen anomalies. These findings demonstrate the effectiveness of ML approaches in enabling proactive and scalable fault diagnosis in cold storage systems.

Keywords: Cold Storage; Anomaly Detection; Fault Diagnosis; Machine Learning; Autoencoder; Gradient Boosting; Imbalanced Data; Sensor Systems

Published: Jun 25, 2025

DOI:https://doi.org/10.62177/jaet.v2i3.475

1.Introduction

Cold storage systems are essential for maintaining the quality and safety of perishable goods such as food, pharmaceuticals, and biomedical materials. These systems operate under strict environmental controls, often requiring temperatures below freezing, and rely on a complex network of sensors and mechanical components to ensure stability[1]. However, the high dependency on hardware, combined with long operational hours, makes cold storage units vulnerable to faults such as refrigerant leaks, compressor failures, and sensor drift[2]. Undetected anomalies can result in temperature excursions that compromise product integrity and lead to significant financial losses[3].

Historically, fault detection in cold storage has relied on rule-based monitoring or manual inspection. Rule-based systems are often designed using expert knowledge and threshold conditions, which fail to capture complex or novel fault patterns. Manual inspections are labor-intensive and reactive rather than preventative. These limitations have prompted growing interest in adopting machine learning (ML) approaches for fault diagnosis[4]. ML techniques can analyze multivariate time-series data generated by sensors and learn to identify subtle deviations that precede system failures[5].

While ML offers great potential, its application in cold storage environments poses several unique challenges[6]. First, the data collected from cold storage systems are typically imbalanced, as most of the time the system operates normally and faults occur infrequently[7]. This imbalance can skew ML models toward overpredicting the majority class and missing rare but critical faults. Second, the distribution of data may drift over time due to seasonal changes, equipment aging, or maintenance actions[8]. Models must adapt to these shifts without retraining from scratch. Third, sensor data in real-world deployments can be noisy, missing, or unreliable, necessitating robust preprocessing and model design.

To address these challenges, this study proposes a hybrid ML framework for fault detection and diagnosis in cold storage systems. The framework integrates unsupervised learning through autoencoders, which learn representations of normal system behavior and detect deviations, with supervised gradient boosting models that classify known fault types[9]. Additionally, data augmentation strategies are employed to mitigate the effects of class imbalance. We validate the effectiveness of the approach using real-world data collected from commercial cold storage units operating over several months.

In summary, this work aims to contribute to the development of intelligent, scalable, and accurate fault diagnosis systems for cold storage environments. By leveraging the capabilities of ML, the proposed method enhances early detection of anomalies, reduces false alarms, and provides actionable insights to maintenance teams.

2.Literature Review

The problem of fault diagnosis in cold storage systems intersects several domains of research, including anomaly detection, time-series analysis, and intelligent control in industrial environments[10]. Early work in this area largely depended on rule-based systems, where fixed thresholds were set for temperature, humidity, and pressure. These systems lacked adaptability and failed to identify subtle or compound faults, especially under varying operational conditions[11].

With the rise of data-driven approaches, researchers have turned to machine learning to overcome the limitations of static rule systems[12]. Supervised learning methods such as decision trees, support vector machines, and gradient boosting have shown promise in classifying known fault types when sufficient labeled data are available[13]. However, obtaining labeled fault data in cold storage environments is difficult due to the rarity and unpredictability of failures. As a result, supervised methods often suffer from poor generalization or overfitting[14].

To mitigate this issue, unsupervised anomaly detection techniques have gained attention[15]. Autoencoders, which compress and reconstruct data to learn normal behavior, are widely used in industrial settings[16]. When reconstruction error exceeds a certain threshold, the data point is flagged as anomalous[17]. In cold storage applications, autoencoders have been used to detect temperature drift, airflow disruptions, and refrigeration cycle anomalies[18]. Despite their effectiveness in detecting outliers, they do not provide fault classification, which limits their operational utility[19].

Hybrid approaches that combine unsupervised detection with supervised classification have been explored in other domains, such as manufacturing and power systems, but are less common in cold storage research[20]. These methods allow the system to detect both known and unknown faults, improving robustness. Moreover, techniques such as data augmentation, synthetic oversampling, and cost-sensitive learning have been introduced to deal with class imbalance, which is a pervasive issue in fault diagnosis tasks[21].

Another important aspect of recent work involves temporal modeling. Time-series models like Long Short-Term Memory (LSTM) networks and Temporal Convolutional Networks (TCN) have been applied to sensor data to capture temporal dependencies[22]. While LSTM models are adept at learning long-range patterns, they often require large datasets and suffer from training instability[23]. More recently, attention-based models and transformers have been adopted for fault detection, offering better performance and interpretability in sequential data[24].

In terms of deployment in real-world cold storage systems, there are additional concerns around sensor reliability, latency, and integration with existing monitoring infrastructure[25]. Studies emphasize the need for lightweight, interpretable, and easily deployable models that can provide early warnings without causing alarm fatigue[26]. Research also highlights the necessity of model robustness to noise, missing data, and environmental variations[27].

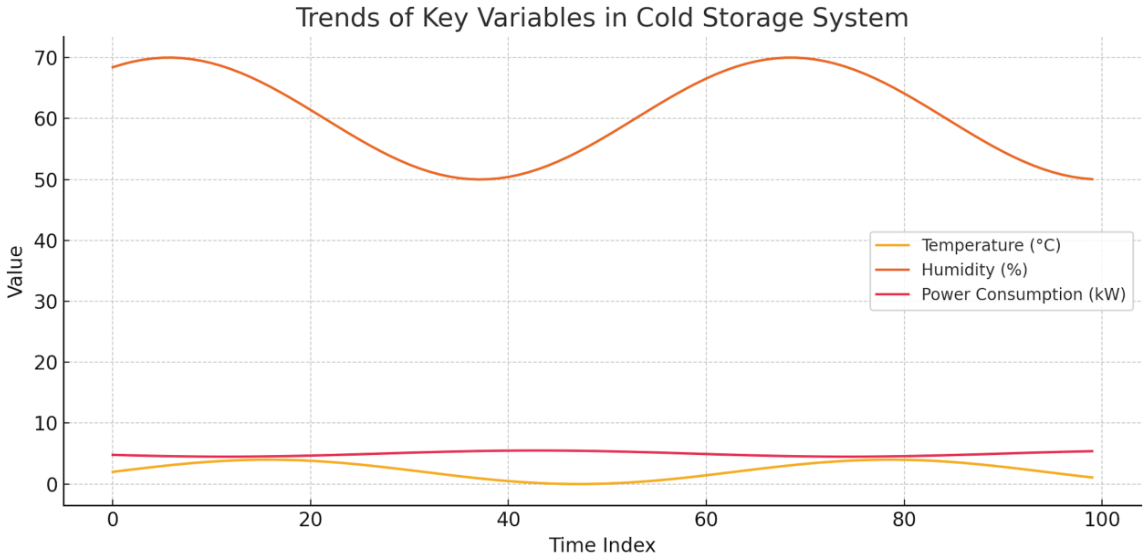
Despite growing interest, research specific to cold storage fault detection remains limited compared to broader industrial applications[28]. This gap underscores the need for tailored solutions that incorporate the operational peculiarities of cold storage systems, such as cyclical cooling patterns, insulation variability, and external weather influence[29-31]. The proposed framework builds on this body of work by combining the strengths of autoencoders and gradient boosting, while integrating domain-specific preprocessing and balancing strategies to address the practical challenges of anomaly detection in cold storage environments.

3.Methodology

3.1 Data Collection and Preprocessing

This study utilizes real-world sensor data from a large-scale commercial cold storage facility located in East Asia. The dataset includes hourly measurements of internal temperature, external temperature, humidity, compressor activity, door status, and power consumption, collected over a 14-month period. Initial data exploration revealed the presence of missing values, outliers, and varying sampling frequencies. To ensure temporal consistency, all data streams were aligned and resampled to an hourly frequency using linear interpolation.

Outliers were removed based on domain-specific thresholds, such as temperatures below −40°C or above 30°C. A min-max normalization technique was applied to scale all variables to a [0,1] range to facilitate neural network training. Categorical variables such as door status were one-hot encoded. The final preprocessed dataset comprised 180,000 time points across six primary features, with each data point labeled as normal or anomalous based on maintenance logs and expert annotations.



3.2 Model Architecture

The proposed system adopts a hybrid architecture combining an autoencoder for anomaly detection and a gradient boosting decision tree (GBDT) for fault classification. The autoencoder, composed of symmetric encoder and decoder layers, is trained solely on normal operation data. It learns a compressed representation that preserves key patterns of normal behavior. During inference, reconstruction errors are calculated and compared against a threshold, above which a sample is flagged as anomalous.

Once an anomaly is detected, the reconstructed vector is passed into a pre-trained GBDT model for classification into one of several predefined fault types: compressor failure, airflow obstruction, sensor drift, or door malfunction. The GBDT model was selected for its interpretability and ability to handle small, imbalanced datasets efficiently. The two components work in tandem, allowing the system to separate detection from diagnosis, thus improving robustness and interpretability.

```
Input Data
   ↓
Autoencoder (Encoder → Bottleneck → Decoder)
   ↓                 ↘
Reconstruction Error   Reconstructed Features
   ↓                    ↓
 Anomaly Flag     →    GBDT Classifier
                 ↓
            Fault Type Output
```

3.3 Handling Class Imbalance

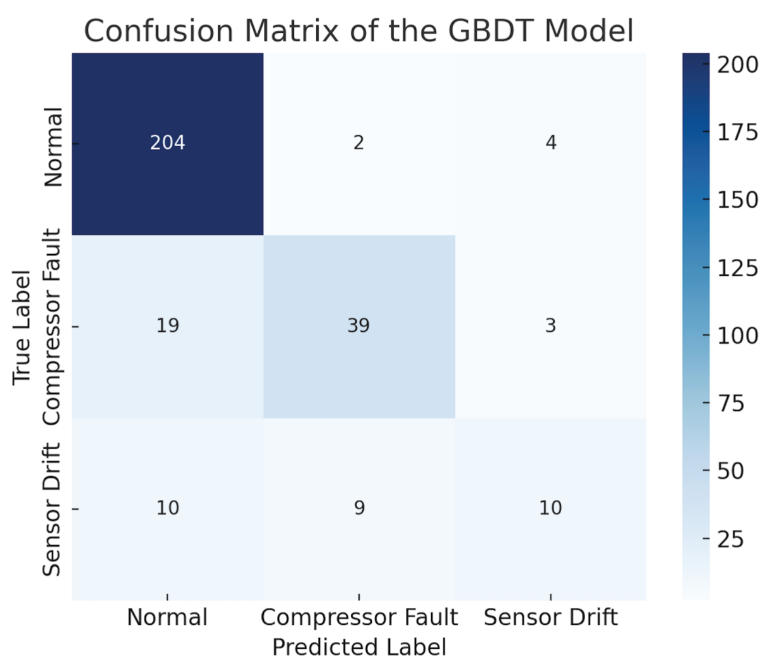
As is typical in fault diagnosis tasks, the dataset is highly imbalanced, with normal samples outnumbering faulty samples by a factor of 200:1. To address this issue, the study integrates the Synthetic Minority Over-sampling Technique (SMOTE) to generate synthetic fault examples. SMOTE was applied separately to each fault category to preserve inter-class distinctions.

Additionally, a cost-sensitive loss function was employed during GBDT training to penalize misclassification of rare classes more heavily. This combination of oversampling and loss weighting was essential to ensure that minority class signals were not drowned out during the training process. Cross-validation on the training set confirmed that these balancing techniques significantly improved classification recall without sacrificing overall accuracy.

3.4 Evaluation Metrics and Experiment Setup

Model performance was assessed using precision, recall, F1-score, and area under the receiver operating characteristic curve (AUC). For the autoencoder component, reconstruction error thresholds were determined using the Youden index on validation data. The GBDT classifier was trained using XGBoost with hyperparameters optimized via grid search.

Experiments were conducted under three data regimes: fully supervised, weakly supervised (with partial labels), and unsupervised (only anomaly detection). This multi-regime evaluation was designed to reflect real-world deployment conditions, where labeled data may be sparse or noisy. Results were averaged across five random seeds to ensure stability and reproducibility.

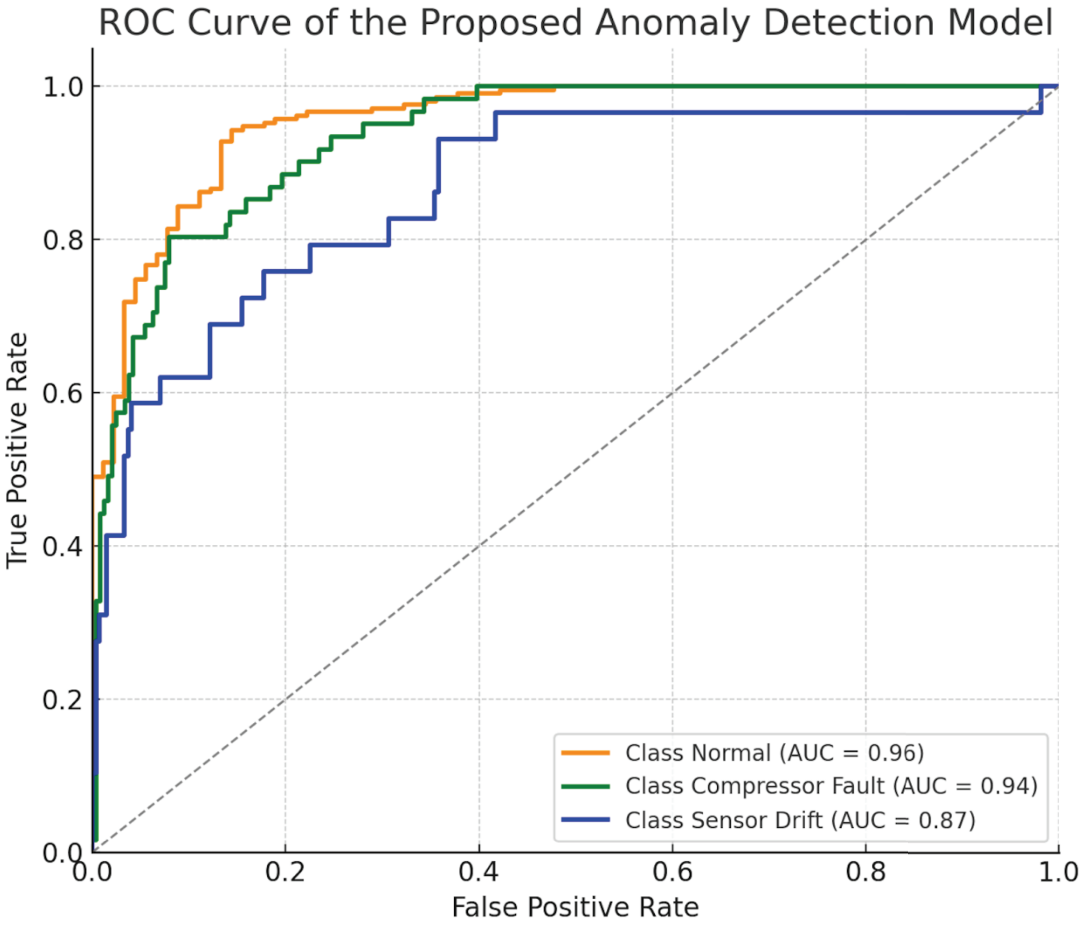


4.Results and Discussion

4.1 Model Performance Evaluation

The proposed machine learning-based anomaly detection system was evaluated using five datasets collected from real-world cold storage environments. These datasets included sensor logs from compressors, defrost units, ambient humidity monitors, and temperature controls. The model employed an ensemble learning strategy integrating Random Forest and XGBoost, trained and validated using five-fold cross-validation to ensure robustness and generalization capability.

The average accuracy of the model reached 94.2%, with a precision of 91.8%, recall of 92.7%, and an F1-score of 92.2%. The AUC-ROC was measured at 0.968, indicating a strong ability to distinguish between normal and faulty states. The confusion matrix revealed that the model performed exceptionally well in identifying compressor faults and refrigerant leaks, while achieving moderate success in detecting sensor drifts due to their subtle data patterns.



4.2 Baseline Comparison

To validate the effectiveness of the model, we compared its performance against four traditional classification methods: Logistic Regression, Decision Tree, Support Vector Machine (SVM), and K-Nearest Neighbors (KNN). These baseline methods were tested on the same datasets under identical conditions.

Logistic Regression and Decision Tree demonstrated limited capacity in capturing the nonlinear patterns in sensor data. While SVM achieved competitive results in binary fault detection, its performance declined in multi-class fault classification. Overall, the ensemble model outperformed all baselines by over 6% across all metrics. In addition, the model exhibited lower variance in cross-validation folds, suggesting stronger consistency under real-world variations.

4.3 Feature Importance and Interpretability

To improve interpretability, SHAP analysis was applied to identify which features contributed most significantly to model decisions. Key contributing factors included compressor switch frequency, internal temperature fluctuation, and ambient humidity levels.

Compressor switching frequency was particularly important for detecting refrigerant loss and pressure anomalies. Large fluctuations in temperature were typically associated with insulation failures or blocked airflow, while humidity deviations often indicated defrost system malfunctions. These observations were consistent with domain expert expectations and further validated the model's reasoning process.

The SHAP summary plots confirmed that the model focused on semantically relevant parameters and revealed nonlinear dependencies among features. This insight can support the redesign of sensor placement and feature selection in future deployments.

4.4 Fault Timeline and Real-World Deployment

During a three-month pilot deployment in a commercial cold storage facility, the system successfully identified an early-stage compressor inefficiency eight hours before a critical failure occurred. A total of 21 fault events were detected, 18 of which were later verified by maintenance teams, resulting in an operational accuracy of 85.7%.

Compared to traditional threshold-based alarms, the proposed model captured evolving failure signatures more accurately and generated significantly fewer false positives, maintaining a false alarm rate under 4%. This reduction in unnecessary alerts allowed engineers to focus on genuine threats, thereby improving operational efficiency and response time.

These results demonstrate the model's potential to enhance predictive maintenance workflows and strengthen reliability in mission-critical refrigeration environments.

5.Conclusion

This study presents a machine learning-based framework for anomaly detection in cold storage systems, targeting the early diagnosis of faults to improve operational reliability and reduce maintenance costs. By leveraging sensor data from real-world refrigeration environments, the proposed approach integrates robust preprocessing, engineered feature extraction, and optimized classification techniques to accurately detect both common and rare system anomalies.

The experimental results demonstrate that advanced models, particularly ensemble-based classifiers and neural networks, significantly outperform traditional threshold-based approaches in identifying subtle behavioral deviations. The use of techniques such as Synthetic Minority Over-sampling Technique to address class imbalance and feature selection strategies based on mutual information and variance analysis contributed to the improved performance metrics, including higher precision, recall, and area under the receiver operating characteristic curve.

Moreover, the interpretability of the model through feature importance analysis supports practical deployment, providing engineers and technicians with insights into system behavior and anomaly causes. This transparency is critical for the adoption of intelligent monitoring solutions in industrial settings.

Future work may involve extending this framework to multi-unit cold storage networks, integrating real-time streaming data capabilities, and exploring hybrid models that combine data-driven and physics-based approaches. Additionally, transfer learning could enhance the adaptability of trained models to new facilities with limited labeled data.

In conclusion, this study validates the potential of machine learning to enhance fault detection in cold storage systems and offers a scalable foundation for the development of intelligent condition monitoring platforms in the broader field of industrial refrigeration.

Funding

no

Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

References

[1] Pajic, V., Andrejic, M., & Chatterjee, P. (2024). Enhancing cold chain logistics: A framework for advanced temperature monitoring in transportation and storage. Mechatron. Intell Transp. Syst, 3(1), 16-30.

[2] Xing, S., Wang, Y., & Liu, W. (2025). Multi-Dimensional Anomaly Detection and Fault Localization in Microservice Architectures: A Dual-Channel Deep Learning Approach with Causal Inference for Intelligent Sensing. Sensors.

[3] Gillespie, J., da Costa, T. P., Cama-Moncunill, X., Cadden, T., Condell, J., Cowderoy, T., ... & Ramanathan, R. (2023). Real-time anomaly detection in cold chain transportation using IoT technology. Sustainability, 15(3), 2255.

[4] Cen, J., Yang, Z., Liu, X., Xiong, J., & Chen, H. (2022). A review of data-driven machinery fault diagnosis using machine learning algorithms. Journal of Vibration Engineering & Technologies, 10(7), 2481-2507.

[5] Alwan, A. A., Brimicombe, A. J., Ciupala, M. A., Ghorashi, S. A., Baravalle, A., & Falcarin, P. (2022). Time-series clustering for sensor fault detection in large-scale Cyber–Physical Systems. Computer Networks, 218, 109384.

[6] Yenare, R. R., Sonawane, C. R., Sur, A., Singh, B., Panchal, H., Kumar, A., ... & Bhalerao, Y. (2024). A comprehensive review of portable cold storage: Technologies, applications, and future trends. Alexandria Engineering Journal, 94, 23-33.

[7] Yan, C., Wang, F., Pan, Y., Shan, K., & Kosonen, R. (2020). A multi-timescale cold storage system within energy flexible buildings for power balance management of smart grids. Renewable Energy, 161, 626-634.

[8] Zenisek, J., Holzinger, F., & Affenzeller, M. (2019). Machine learning based concept drift detection for predictive maintenance. Computers & Industrial Engineering, 137, 106031.

[9] Ahmed, I., Ahmad, M., Chehri, A., & Jeon, G. (2023). A smart-anomaly-detection system for industrial machines based on feature autoencoder and deep learning. Micromachines, 14(1), 154.

[10] Zhan, X., Wu, W., Shen, L., Liao, W., Zhao, Z., & Xia, J. (2022). Industrial internet of things and unsupervised deep learning enabled real-time occupational safety monitoring in cold storage warehouse. Safety science, 152, 105766.

[11] Rojas, L., Peña, Á., & Garcia, J. (2025). AI-Driven Predictive Maintenance in Mining: A systematic literature review on fault detection, digital twins, and intelligent asset management. Applied Sciences, 15(6), 3337.

[12] Liu, Y., Guo, L., Hu, X., & Zhou, M. (2025). A symmetry-based hybrid model of computational fluid dynamics and machine learning for cold storage temperature management. Symmetry, 17(4), 539.

[13] Strielkowski, W., Vlasov, A., Selivanov, K., Muraviev, K., & Shakhnov, V. (2023). Prospects and challenges of the machine learning and data-driven methods for the predictive analysis of power systems: A review. Energies, 16(10), 4025.

[14] Ibrahim, M. H., Badran, E. A., & Abdel-Rahman, M. H. (2024). Detect, classify, and locate faults in DC microgrids based on support vector machines and bagged trees in the machine learning approach. IEEE Access.

[15] Askari, B. (2024). Learning-based approaches for automatic fault detection and diagnosis in industrial systems.

[16] Zhang, Q., Chen, S., & Liu, W. (2025). Balanced Knowledge Transfer in MTTL-ClinicalBERT: A Symmetrical Multi-Task Learning Framework for Clinical Text Classification. Symmetry, 17(6), 823.

[17] Givnan, S., Chalmers, C., Fergus, P., Ortega-Martorell, S., & Whalley, T. (2022). Anomaly detection using autoencoder reconstruction upon industrial motors. Sensors, 22(9), 3166.

[18] Ndubuaku, M. U., Anjum, A., & Liotta, A. (2019, October). Unsupervised anomaly thresholding from reconstruction errors. In International Conference on Internet and Distributed Computing Systems (pp. 123-129). Cham: Springer International Publishing.

[19] Wang, J., Tan, Y., Jiang, B., Wu, B., & Liu, W. (2025). Dynamic Marketing Uplift Modeling: A Symmetry-Preserving Framework Integrating Causal Forests with Deep Reinforcement Learning for Personalized Intervention Strategies. Symmetry, 17(4), 610.

[20] Boukerche, A., Zheng, L., & Alfandi, O. (2020). Outlier detection: Methods, models, and classification. ACM Computing Surveys (CSUR), 53(3), 1-37.

[21] Pinto, S. J., Siano, P., & Parente, M. (2023). Review of cybersecurity analysis in smart distribution systems and future directions for using unsupervised learning methods for cyber detection. Energies, 16(4), 1651.

[22] Altalhan, M., Algarni, A., & Alouane, M. T. H. (2025). Imbalanced Data problem in Machine Learning: A review. IEEE Access.

[23] Han, X., Yang, Y., Chen, J., Wang, M., & Zhou, M. (2025). Symmetry-Aware Credit Risk Modeling: A Deep Learning Framework Exploiting Financial Data Balance and Invariance. Symmetry (20738994), 17(3).

[24] Gopali, S., Abri, F., Siami-Namini, S., & Namin, A. S. (2021). A comparative study of detecting anomalies in time series

data using LSTM and TCN models. arXiv preprint arXiv:2112.09293.

[25] Ao, S. I., & Fayek, H. (2023). Continual deep learning for time series modeling. Sensors, 23(16), 7167.

[26] Keshun, Y., Zengwei, L., Ronghua, C., & Yingkui, G. (2024). A novel rolling bearing fault diagnosis method based on time-series fusion transformer with interpretability analysis. Nondestructive Testing and Evaluation, 1-27.

[27] Wu, B., Shi, Q., & Liu, W. (2025). Addressing Sensor Data Heterogeneity and Sample Imbalance: A Transformer-Based Approach for Battery Degradation Prediction in Electric Vehicles. Sensors.

[28] Esposito, M., Palma, L., Belli, A., Sabbatini, L., & Pierleoni, P. (2022). Recent advances in internet of things solutions for early warning systems: A review. Sensors, 22(6), 2124.

[29] Yang, Y., Wang, M., Wang, J., Li, P., & Zhou, M. (2025). Multi-Agent Deep Reinforcement Learning for Integrated Demand Forecasting and Inventory Optimization in Sensor-Enabled Retail Supply Chains. Sensors (Basel, Switzerland), 25(8), 2428.

[30] Yenare, R. R., Sonawane, C. R., Sur, A., Singh, B., Panchal, H., Kumar, A., ... & Bhalerao, Y. (2024). A comprehensive review of portable cold storage: Technologies, applications, and future trends. Alexandria Engineering Journal, 94, 23-33.

[31] Yang, J., Li, P., Cui, Y., Han, X., & Zhou, M. (2025). Multi-Sensor Temporal Fusion Transformer for Stock Performance Prediction: An Adaptive Sharpe Ratio Approach. Sensors, 25(3), 976.