Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Ensemble-Based Anomaly Detection for Urban Power Efficiency Towards Sustainable Energy Management

Visalakshi Annepu
VIT-AP University
Kalapraveen Bagadi
VIT-AP University
Muhsin Jaber Jweeg
Al-Farahidi University
Mohd Shukor Salleh
Fakulti Teknologi Dan Kejuruteraan Industri Dan Pembuatan, Universiti Teknikal Malaysia Melaka, Hang Tuah Jaya
Deepika Rani Sona
Vellore Institute of Technology
Oday I. Abdullah
oday.abdullah@coeng.uobaghdad.edu.iq
University of Baghdad
Oula M. H. Fatla
Gulf University
Munaf S. Majeed
Al-Nahrain University

Article

Keywords: Energy Efficiency, Anomaly Detection, Smart Buildings, Power Consumption, Machine Learning

Posted Date: October 7th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11097255/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Ensemble-Based Anomaly Detection for Urban Power Efficiency Towards Sustainable Energy Management

Visalakshi Annepu[1], Kalapraveen Bagadi[2], Muhsin Jaber Jweeg[3], Mohd Shukor Salleh[4], Deepika Rani Sona[5], Oday I. Abdullah[6,7], Oula M. H. Fatla[8], Munaf S. Majeed[9,10]

[1]School of Computer Science and Engineering, VIT-AP University, Amaravati 522 241, India
annepuvisala@gmail.com
[2]School of Electronics Engineering, VIT-AP University, Amaravati 522 241, India
kpbagadi@gmail.com
[3]College of Technical Engineering, Al-Farahidi University, Baghdad 00965, Iraq
muhsin.jweeg@uoalfarahidi.edu.iq
[4]Fakulti Teknologi Dan Kejuruteraan Industri Dan Pembuatan, Universiti Teknikal Malaysia Melaka, Hang Tuah Jaya, 76100 Melaka, Malaysia;
shukor@utem.edu.my
[5]School of Electronics Engineering, Vellore Institute of Technology, Vellore 632 014, India
deepika.rs@vit.ac.in
[6]Department of Energy Engineering, College of Engineering, University of Baghdad, Baghdad, Iraq
[7]College of Engineering, Al-Naji University, Baghdad, Iraq
oday.abdullah@coeng.uobaghdad.edu.iq
[8]Mechanical Engineering Department, College of Engineering, Gulf University, Sanad 26489, Bahrain
dr.oula.fatla@gulfuniversity.edu.bh
[9]AL-Nahrain Renewable Energy Research Center, Al-Nahrain University, Baghdad, Iraq
[10]Department of Radiology Techniques, Dijlah University College, Baghdad, Iraq
munafsm@nahrainuniv.edu.iq

Corresponding Author: oday.abdullah@coeng.uobaghdad.edu.iq

Ensemble-Based Anomaly Detection for Urban Power Efficiency Towards Sustainable Energy Management

Abstract: To address growing sustainability objectives and urban energy requirements, this research developed an anomaly detection system for urban infrastructure energy efficiency. For anomaly detection in energy consumption, the study evaluates traditional machine learning algorithms, such as Isolation Forest, Local Outlier Factor (LOF), and One-Class SVM, in addition to deep learning Autoencoders. While missing value treatment, feature encoding, and scaling for data training are necessary for model effectiveness, the study applied feature engineering to enhance model performance. Evaluation metrics include precision, recall, F1 score, and anomaly detection rate. Based on experimental simulation of ground truth data, the Autoencoder most distinguished itself with F1 metrics of 0.9740 and 0.9494 in overall accuracy. The application of the SHapley Additive exPlanation (SHAP) framework on focus features to explain model predictions and the overall system energy efficiency and conservation impact reinforces the framework's outcome with confidence. Enhancements to this work include time-series anomaly detection and dynamic thresholds for real-time detection.

Keywords: Energy Efficiency, Anomaly Detection, Smart Buildings, Power Consumption, Machine Learning.

1. Introduction

The swift transformation of portions of a population into urban centers requires effective and sustainable solutions in energy management to maintain civic infrastructure. For civic societies, which comprise residential, commercial, and public facilities, the provision of uninterrupted and secure power systems is a necessity in maintaining operational functionality and the quality of life of the citizens [1]. The implementation of sustainable initiatives becomes a challenge, as the energy inefficiencies are hidden within unscheduled system spikes and breakdowns, as well as other inefficient uses of energy [2]. The unaddressed power inefficiencies operationally cost money and waste/destroy the potential to uplift the lost energy, while also increasing the environmental footprint, thus demanding more sophisticated means to detect and eliminate inefficiencies [3]. Civic energy inefficiencies are in dire need of management, as the European Commission cites buildings as responsible for 40% of energy use and 36% of energy-related greenhouse gas emissions in the European Union [4].

The use of data in energy management systems is enabled by the application of anomaly detection systems, which are essential for identifying illegal power usage and flows [5]. The New energy datasets are more complex and multidimensional that they outperform the traditional statistical data set [6]. It is the rapid volume and variability of data that drives the need for more advanced analytical solutions to be put in place to detect, in a more effective manner, energy system inefficiencies [7].

The progress of machine learning (ML) and deep learning (DL) methods in the field has provided the ability to develop efficient tools for anomaly detection in the identification of intricate patterns [8]. Studies show the Isolation Forest's ability to detect outliers in large datasets, while Local Outlier Factor performs density-based anomaly detection, and One-Class Support Vector

Machines detect normal and abnormal patterns without supervision [9, 10, 11]. Autoencoders, because of their advanced features, can extract intricate and complex patterns in the data, making them suitable for irregular behaviors in the consumption of energy [12]. In addition, the use of ensemble methods to enhance predictive robustness by integrating different model outputs is an approach that has been increasingly adopted for use in energy anomaly detection [13].

The proposed research aims to develop a complete anomaly detection system for power consumption data to improve the energy efficiency of civic society. Performance evaluation includes traditional ML models such as Isolation Forest, LOF, One-Class SVM, Deep Learning Autoencoders on a fully pre-processed and engineered dataset. Predictive robustness is increased for ensemble modelling through the integration of several unsupervised ML models with SHAP (SHapley Additive exPlanations) values for explainable AI on feature contributions for clustered anomaly detection [14]. The Autoencoder and Isolation Forest techniques excel at detecting patterns of irregular energy use, thus illustrating potential avenues to be explored for energy savings. The suggested framework is promising for real-time civic monitoring yet primarily aims to facilitate research on time series analytics and adaptive threshold methods for improving sustainable energy management for urban areas [15].

This study's key contributions include the following:

• Comprehensive Preprocessing Pipeline: it describes a formal data preprocessing strategy with the following components: mean imputation, log transformation, feature scaling using MinMax normalization, low-variance feature deletion, and skewness reduction through semi-interquartile range extraction.

• Employing a Civic Energy Dataset, it analyzes and benchmarks various techniques used in Anomaly Detection, comparing standard machine learning approaches (Isolation Forest, Local Outlier Factor, and One-Class SVM) to a deep learning-based Autoencoder in identifying abnormal consumption patterns.

• Using a majority voting approach, the Ensemble-Based Detection Framework integrates various model components to enhance system robustness and decrease the prevalence of false positives in civic energy monitoring.

• SHAP (SHapley Additive exPlanations) analyzes and interprets the assisting elements used in the prediction of anomalous behavior and, as such, aids civic administrators when deciding on applicable energy conservation measures.

The paper is organized into five main sections. Section 1 introduces the motivation for sustainable energy management and highlights the need for anomaly detection in urban power systems. Section 2 reviews related research on machine learning and deep learning methods for energy anomaly detection. Section 3 details the proposed framework, algorithms, and preprocessing steps. Section 4 presents performance evaluation, visualization, and interpretability analysis. Section 5 concludes with key findings, limitations, and future research directions for real-world implementation.

2. Literature Review

Urban energy consumption efficiency is a focal point of concern and research, as urbanization is developing at a fast pace and energy consumption is growing to unprecedented levels [16]. Anomaly detection is an important analytical technique used for identifying patterns and irregularities in energy consumption for utilities [17]. The proposed framework is built upon a

literature review of energy management, anomalies detection techniques, and the impact of civic infrastructure on urban societies.

Historically, statistical methods were the primary approach used in energy systems for anomaly detection. Z-score methods and threshold limits were applied for the analyses to assess deviations in reported data and identify power use anomalies [18]. Such techniques were found to be ineffective on real-world energy datasets, as they did not capture the combination of variables that involve civic environment patterns, including occupancy, equipment behavior, and weather [19]. Lack of proficiency in traditional methods catalyzed the shift toward adopting machine learning (ML) and deep learning (DL) frameworks, as they provide accurate, adaptable techniques for solving the problems at hand [20].

The Isolation Forest algorithm is a widely used algorithm in ML because it quickly identifies anomalies in high-dimensional data sets [21]. This method, in previous works, was established as a data partitioning method that scales energy monitoring systems by partitioning data to control anomaly detection. It has been tested in solar power plants to identify equipment failures and abnormal use patterns, which is highly relevant for civic energy systems [22]. The Local Outlier Factor (LOF) anomaly detection method uses the local density of data points to locate energy consumption anomalies effectively [23]. The method has been tested on building energy analysis, but it was mentioned that much of its performance relies on closely controlled settings and adjustments around its parameters, and particularly the spatial scale of the local-ness [24].

There has been a growing interest in conducting anomaly detection without oversight using One-Class Support Vector Machines (One-Class SVM) [25]. This approach improves the detection of boundaries that distinguish regular data samples from irregular data and is optimal for smaller datasets, which aligns with the requirements of civic energy applications. For the research, the methodology was proven effective in consumption data for meter robust fault detection and for detecting spontaneous demand spikes [26]. The described approach loses its power with high-dimensional data, as investigated in this research, until scholars find a way to merge the approach with suitable data feature selection.

The use of deep learning techniques, especially autoencoders, allows organizations to detect anomalies through neural networks that identify complex structural patterns within data [27]. Autoencoders demonstrate a data reconstruction functionality that presents anomalies through high error rates. They showed the capability to detect energy consumption patterns in smart grids based on reviews of time-series anomaly detection methods [28]. Autoencoder systems protect civic power distribution networks by identifying weak anomalies through implementations that outperform traditional ML models [29].

The literature uses ensemble methods as a predictive model performance improvement strategy through model aggregation [30]. Research explained ensemble methods as a way to reduce model bias while making systems more robust. Anomaly detection methods are linked through ensemble techniques in the energy systems field, as noted in papers on industrial energy system fault detection that joined SVM with Isolation Forest detection [31]. The research applies ensemble modelling to integrate the anomaly detection features of multiple models.

Adopting civic applications for the use of anomaly detection models comes with certain interpretability needs that stem from considerable unease. The SHAP methodology provides model justification by explaining feature attribution [32]. SHAP is used in research with ML to identify anomalies in building energy data and their associated inefficiencies. Civic stakeholders

implement energy-saving actions, which is why this method is particularly relevant in these contexts, as they need straightforward, actionable recommendations [33].

Although these studies suggest accomplishments, there are still outstanding issues. Much of the literature tends to concentrate on particular domains, such as industrial or renewable energy, leaving the wider civic mosaic that includes the domestic, commercial, and civic spheres, largely unaddressed. The civic energy research remains largely unexamined concerning the use of preprocessing techniques and their impact on model performance. This work surveys the landscape by conducting knowledge model benchmarking of machine and deep learning algorithms on processed data and implementing ensemble system evaluation and SHAP interpretability to optimize power efficiency within civic edifices.

3. Methodology

This research develops a strong anomaly detection system that identifies irregularities within power consumption data to maximize civic energy efficiency. The implementation process comprises data collection and cleaning as well as model selection and training, followed by ensemble system integration and performance evaluation and interpretation capabilities. The system's steps were created specifically to handle civic energy data intricacies for execution in urban monitoring systems. The project makes use of the Python programming language with pandas, scikit-learn, tensorflow, numpy, shap, and matplotlib libraries for its execution.

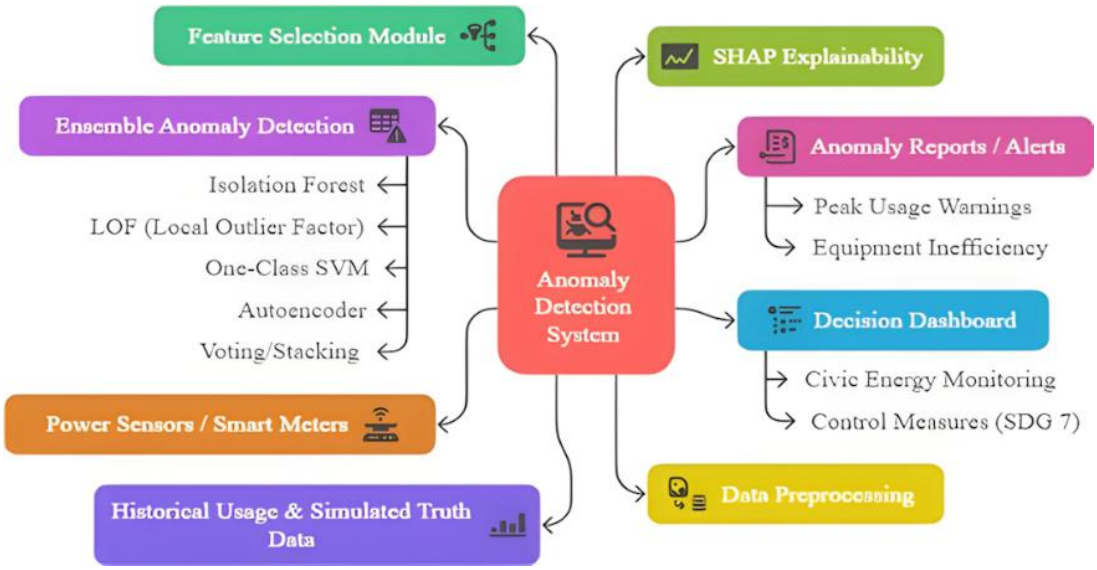


Figure 1: Architecture of the Energy Monitoring Anomaly Detection System.

Figure 1 shows that the energy monitoring anomaly detection system architecture demonstrates the ability to scale urban user energy consumption tracking while maintaining interpretation capabilities and accuracy for aberrant trend detection. Smart meters and IoT devices supply Civic Energy Data that initiates the system process. Before processing, the raw data undergoes normalization, imputation, and modification steps. The Feature Engineering module receives processed data to create suitable statistical and time-series characteristics for performance enhancement. The main Processing Layer combines four machine learning models, including One-Class SVM, Local Outlier Factor (LOF), Isolation Forest, and Autoencoder, which analyse data anomalies through separate detection perspectives. ADTL moves data to its Decision Layer,

which combines the Ensemble Voting method to produce more resilient outputs, whereas SHAP Explainability analyses feature influence on the anomaly score. Results / Alerts generates stakeholder-relevant alerts through the process that accepts anomaly ratings and explanations from the final insight module. Model lifecycle control and process repeatability happen through the use of MLflow, which organizes the complete operation.

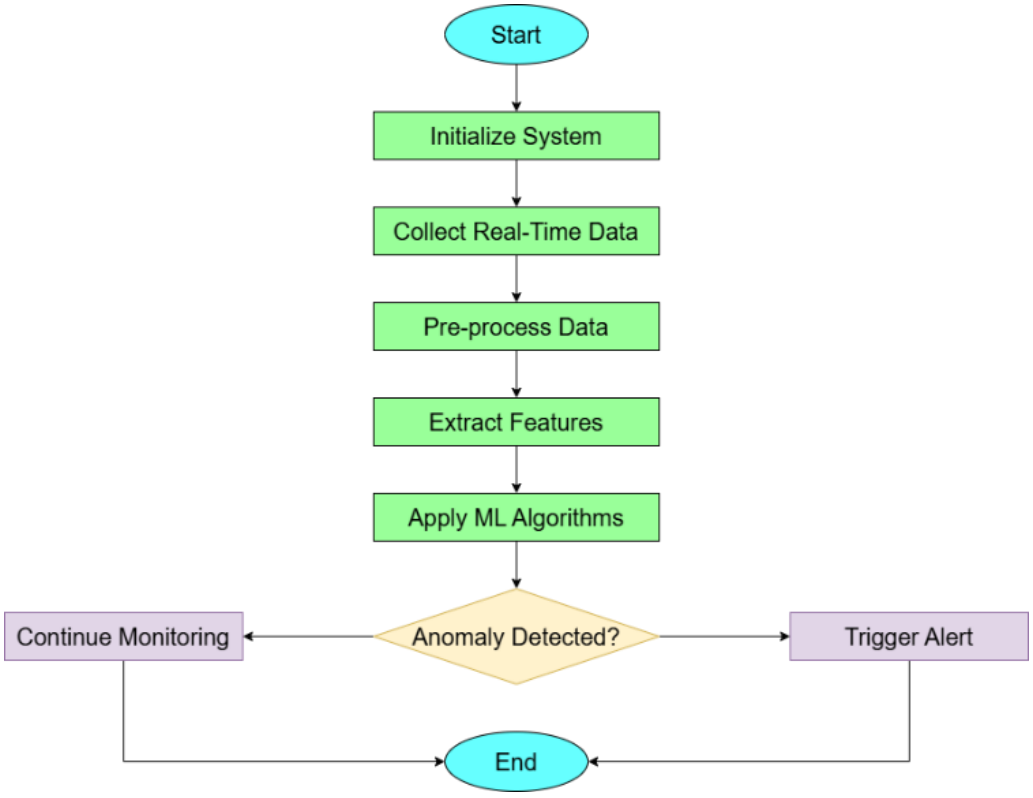


Figure 2: A flowchart of the proposed model.

The designed flowchart in Figure 2 shows the step-by-step functioning of the anomaly detection system used for real-time energy monitoring. The system commences with the Start node and proceeds to System Initialization, which enables all required components and functionalities. The system receives current sensor and smart meter data through its processes. Preprocessing steps, which include cleaning, normalization, and formatting, should be applied to raw data before the machine learning pipeline can use it effectively. Beneficial data characteristics appear through Feature Extraction after processing the pre-processed input data. The Forecasting Tool delivers data points to the Machine Learning Model to allow analysis that identifies atypical patterns during dataset reviews. If the Model detects such activity at the Decision Node, then the system performs stakeholder notification. The system checks for anomalies. In case of no detected abnormalities, the process moves back to the monitoring phase and gathers new data. The loop maintains a continuous and autonomous operation of the system. The system transitions to the End node either when the monitoring period terminates, or the system shutdown occurs.

3.1 Data Acquisition and Preprocessing

The power consumption records in the dataset originate from municipal operations, with particular aspects (source and time duration) omitted for this analysis. R1.1: The timestamp column is removed during initial preprocessing since static features assessment requires exclusivity. The

SimpleImputer from scikit-learn performs mean imputation to complete the data through its implementation for missing value handling. The encoding of categorical features takes a numerical form as per standard practice in civic energy data preparation, although the dataset does not explicitly mention this encoding method. Missing values are imputed using mean imputation as in Equation (1):

$$y_i' = \begin{cases} y_i, & \text{if } y_i \text{ is not missing} \\ \frac{1}{m}\sum_{i=1}^{m} y_j, & \text{if } y_i \text{ is missing} \end{cases} \quad (1)$$

where y_i' is the imputed value of the i-th feature after mean imputation, y_i is the original value of the i-th feature in a sample, y_j are the non-missing values of the feature across all samples, and m is the number of non-missing samples for the feature. To address skewness, a logarithmic transformation is applied as given below:

$$y_i'' = \log(1 + y_i') \quad (2)$$

where, y_i' is the imputed feature value before transformation, and y_i'' is the transformed feature value after applying $\log(1 + y_i')$. The model implements Function Transformer together with (np.log1p) for skewness correction in energy consumption data without altering the non-negative properties and enhancing variance stability. Through normalization by MinMaxScaler, the features end up ranging from 0 to 1 while helping models maintain stable performance in cases of variable scales, as given below:

$$y_i''' = \frac{y_i'' - \min(y'')}{\max(y'') - \min(y'')} \quad (3)$$

where y_i'' is the transformed feature value post-log transformation, and y_i''' is the scaled feature value in [0, 1]; $\min(y'')$ and $\max(y'')$ are the minimum and maximum values of the feature after log transformation across all samples. The VarianceThreshold function with a threshold of 0.01 removes uninformative low-variance features, which decreases both computational workload and dimensional scale yet preserves important attribute information. The preprocessed data set is divided into training sections that represent 80% and testing sections 20% of the total using train_test_split and random state 42 for replicability.

$$Var(y) = \frac{1}{m}\sum_{i=1}^{m}(y_i - \mu)^2 \quad (4)$$

where y_i is the value of the i-th sample for a feature, μ is the mean of the feature across all samples $\left(\mu = \frac{1}{m}\sum_{i=1}^{m} y_i\right)$, m is the total number of samples, and $Var(y)$ is the variance of the feature, compared to 0.01.

3.2 Model Selection and Training

The selected anomaly detection models belong to four proven systems that harness previous research findings and support unsupervised learning due to the absence of labelled anomalies in civic energy data.

Isolation Forest: When utilizing the IsolationForest from the scikit-learn library, it use a contamination rate of 0.1 to indicate 10% anomalies and sets a random state value to 42. Random data partitioning allows this approach to identify outliers through its high efficiency with large-scale multivariable datasets.

Local Outlier Factor: The Local Outlier Factor (LOF) model with parameters 20 neighbours and 0.1 contamination identifies anomalies by studying local density variations, thus being sensitive to contextual irregularities.

One-Class Support Vector Machines: It is used with an RBF kernel, and a μ value of 0.1 allows the model to create a separating boundary between regular data and outliers when operating in situations with minimal labelled data.

Autoencoder: This TensorFlowkeras model implements an autoencoder design with layers consisting of 16 input, 8 bottleneck, and 16, then 16 output units set to match the input dimensions. The model contains a hidden layer with ReLU activation and output sigmoid activation, trained for 50 epochs with the Adam optimizer, minimizing mean squared error (MSE) in batches of 16 samples. To identify anomalies, the reconstruction error is thresholded at the 95th percentile level. The pre-processed X_train data is subjected to model training, which uses the autoencoder function to detect normal patterns in the training data. The X_test predictions receive a binary transformation that changes negative-one anomalies to zero values and normal observations to ones through the np.where function.

3.3 Ensemble Modelling

A combination approach called ensemble utilizes majority voting procedures to integrate outputs from these four predictive models. This algorithm creates predictions through np.vstack before classifying a sample as anomalous (0) when at least two models detect it as abnormal; otherwise, normal (1), as in Eq. (5). The approach utilizes different detection power strengths between Isolation Forest and LOF to minimize the weaknesses of independent models.

$$Ensemble(y) = \begin{cases} 0, & if \sum_{m=1}^{N} I\left(pred_m(y) = 0\right) \geq 2 \\ 1, & otherwise \end{cases} \tag{5}$$

where $Ensemble(y)$ is the final prediction for sample y (0 or 1), m is the number of models, $pred_m(y)$ is the prediction of the m-th model for sample y (0 or 1), and I is the indicator function (1 if condition is true, 0 otherwise).

3.4 Performance Evaluation

The performance assessment employs accuracy and precision together with recall and F1-score metrics that are computed in sklearn.metrics. The dataset lacks ground truth labels, so the synthetic evaluation operates under the assumption that all test samples are normal, while models need to detect 10% contamination as anomalies. The evaluation process uses an authentic civic scenario that demonstrates unusual cases happening infrequently. Given the data values matrix X_test,

synthetic evaluation generates visual insights through bar charts for metric comparison and scatter plots for anomaly detection, together with box plots for score distributions and TSNE for 2D anomaly visualization and anomaly detection plots using the matplotlib and seaborn packages. The metrics such as accuracy, precision, recall, and F1-score are expressed mathematically as:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{6}$$

$$Precision = \frac{TP}{TP + FP} \tag{7}$$

$$Recall = \frac{TP}{TP + FN} \tag{8}$$

$$F1 = 2\frac{Precision \times Recall}{Precision + Recall} \tag{9}$$

where *TP* is true positives (correctly predicted anomalies), *TN* is true negatives (correctly predicted normals), *FP* is false positives (normals predicted as anomalies), and *FN* is false negatives (anomalies predicted as normal).

3.5 Interpretability Analysis

SHapley Additive exPlanations (SHAP) values are computed through shap.Explainer for the autoencoder to achieve transparency requirements that civic applications need. The summary plot shows results of feature contribution analysis for X_test anomaly predictions performed by this operation. The SHAP given in eq. (10) system enhances stakeholder understanding through feature-specific anomaly-impact calculation, which meets actual energy management needs.

$$\varphi_i = \sum_{S \in N\{i\}} \frac{|S|!(|N| - |S| - 1)!}{|N|!} \left[f(S \cup \{i\}) - f(S) \right] \tag{10}$$

where,φ_i is the SHAP value for feature *i*, *S* is the subset of features excluding *i*, *N* is the total set of features, |*S*| is the size of subset *S*, |*N*| is the total number of features, and *f*(*S*) is the model output with features in *S*.

3.6 Implementation Details

The framework operates on Python 3.x using pandas to handle data, while scikit-learn executes ML models and pre-processing steps and tensor-flow provides the autoencoder functionality, and NumPy performs numerical operations, followed by interpretability through shape and visualizations made possible by matplotlib/seaborn libraries. Future enhancements of time-series or real-time civic energy data will benefit from the workflow design, which promotes scalability and allows for reproducible results. Algorithm 1 cleans and prepares energy data by imputing missing values, applying log transformation, normalizing with Min–Max scaling, and removing low-variance features, producing a reliable dataset for model training.

Algorithm 1. Data Preprocessing and Feature Selection for Anomaly Detection.

Inputs:

- Raw dataset D
- Missing-value tolerance τ
- Variance threshold θ
- Feature set $\{f_1, f_2, \ldots, f_n\}$

Output:

- Pre-processed dataset D_{prep}.

Steps:

1. **Load and sanitize dataset**
   1.1 Import the dataset D.
   1.2 Remove non-informative or purely temporal fields (e.g., Timestamp) to focus on static behavioral/consumption features.
2. **Handle missing data**
   For each feature f_i in D:
   2.1 Compute the fraction of missing values in f_i.
   2.2 If the fraction of missing values in f_i is less than or equal to τ:
   • Replace missing entries in f_i using mean imputation:

   $\hat{f}_i = \frac{1}{m}\sum_{j=1}^{m} f_i^{(j)}$, where m is the number of available (non-missing) samples for that feature.

   2.3 Otherwise (if the feature is mostly missing):
   • Optionally discard f_i from further analysis.
3. **Stabilize skewed distributions**
   For each retained (imputed) feature value v:
   • Apply a log-based transformation to reduce skew while preserving non-negativity: $v' = \log(1 + v)$.
   This step helps compress extreme consumption spikes in civic energy data and improves model robustness.
4. **Scale features to a uniform range**
   For each transformed feature v':
   • Apply Min–Max normalization to map values to $[0,1]$: $v_{scaled} = \frac{v' - \min(v')}{\max(v') - \min(v')}$.
   This prevents high-magnitude features (e.g., heavy-load consumption lines) from dominating distance- or density-based anomaly detectors.
5. **Remove low-variance / uninformative features.**
   For each feature f_i after scaling
   5.1 Estimate its variance:

   $Var(f_i) = \frac{1}{m}\sum_{j=1}^{m}\left(f_i^{(j)} - \mu_i\right)^2 \quad \text{where} \quad \mu_i = \frac{1}{m}\sum_{j=1}^{m} f_i^{(j)}$

   5.2 If $Var(f_i) < \theta$
   • Drop f_i (feature is nearly constant and not useful for anomaly separation).
   5.3 Else
   Keep f_i.
6. **Produce final learning set**
   6.1 Collect all retained, scaled features into D_{prep}.
   6.2 Split D_{prep} into training and testing subsets (e.g., 80% / 20%) with a fixed random state for reproducibility.

Return:

D_{prep}, the cleaned, transformed, scaled, and variance-filtered dataset ready for model training and evaluation.

Algorithm 2 integrates Isolation Forest, LOF, One-Class SVM, and Autoencoder outputs through majority voting, flagging samples as anomalies when at least two models agree.

Algorithm 2. Ensemble Majority-Voting Anomaly Decision

Inputs:

- Test samples $x \in X_{\text{test}}$
- Binary predictions from M anomaly detectors
  $\{p_1(x), p_2(x), \ldots, p_M(x)\}$, where each $p_m(x) \in \{0, 1\}$
  - Convention:
    - 0= "anomalous / abnormal consumption"
    - 1= "normal consumption"
- Voting threshold T (minimum number of models that must agree on 'anomaly')

Output:

- Final ensemble label Ensemble(x) for every test sample x

Steps:

1. **Initialize ensemble parameters**
   1.1 Let $M = 4$ models: Isolation Forest, Local Outlier Factor (LOF), One-Class SVM, and Autoencoder.
   1.2 Set voting threshold $T = 2$.
   (Meaning: at least two independent detectors must agree that a sample is abnormal before escalating it as an anomaly.)
2. **Aggregate model opinions per sample**
   For each test sample $x \in X_{\text{test}}$:
   2.1 Count how many models labeled it as anomalous:
   $$c(x) = \sum_{m=1}^{M} \mathbf{1}[p_m(x) = 0]$$
   Where **1**[.] is the indicator function that returns 1 if the condition is true and 0 otherwise.
3. **Apply majority logic**
   For each x:
   3.1 If $c(x) \geq T$:
   • Set Ensemble(x) = 0 (flag as anomaly / potential energy inefficiency or fault condition).
   3.2 Else:
   • Set Ensemble(x) = 1 (treat as normal behavior).
4. **Return decisions**
   4.1 Output Ensemble(x) for all $x \in X_{\text{test}}$.
   4.2 These ensemble labels are then used for downstream alerting, visualization, and SHAP-based interpretation.

4. Results and Discussions

This section displays an evaluation of Isolation Forest alongside Local Outlier Factor (LOF), One-Class SVM, and Autoencoder, and their ensemble through analysis of power consumption data performance. Quantitative metrics together with visualizations form the basis of results, while links to civic society energy efficiency form the discussion points. Table 1 summarizes the parameters for proposed models.

Table 1: Parameter Settings for Anomaly Detection Models and Preprocessing

Parameter	Value
Pre-processing	
Imputation	Mean (SimpleImputer)
Transformation	Log (np.log1p)

Scaling	MinMax [0,1]
Variance Threshold	0.01
Train – Test Split	80% train, 20% test
Random State	42
Isolation Forest	
Contamination	0.1
Random State	42
Local Outlier Factor	
Contamination	0.1
Random State	42
One-Class SVM	
Kernel	RBF
Nu	0.1
Autoencoder	
Hidden Layers	[16, 8, 16, 16]
Activation (Hidden)	ReLU
Activation (Output)	Sigmoid
Epochs	50
Batch Size	16
Optimizer	Adam
Loss Function	MSE
Anomaly Threshold	95^{th} percentile
Ensemble	
Voting Mechanism	Majority (≥ 2 models)

4.1 Performance Metrics

The testing models analyzed the preprocessed X_test dataset using synthetic ground truth (y_test = np.ones(len(X_test))) related to a 10% contamination rate for anomaly simulation. The research implemented performance metrics: accuracy together with precision and recall, in addition to F1-score, while following the methodology for definitions. Table 2 shows the comparison metrics of all the ML models. Figure 3 shows a Bar chart comparing accuracy, precision, recall, and F1-score across models, with the Autoencoder showing the highest F1-Score (0.9740).

Table 2: Comparison of ML Models with metrics.

Model	Accuracy	Precision	Recall	F1 Score
Isolation Forest	0.8924	1.0000	0.8924	0.9431
Local Outlier Factor (LOF)	0.8987	1.0000	0.8987	0.9467
One-Class SVM	0.8987	1.0000	0.8987	0.9467
Autoencoder	0.9494	1.0000	0.9494	0.9740
Ensemble	0.8861	1.0000	0.8861	0.9396

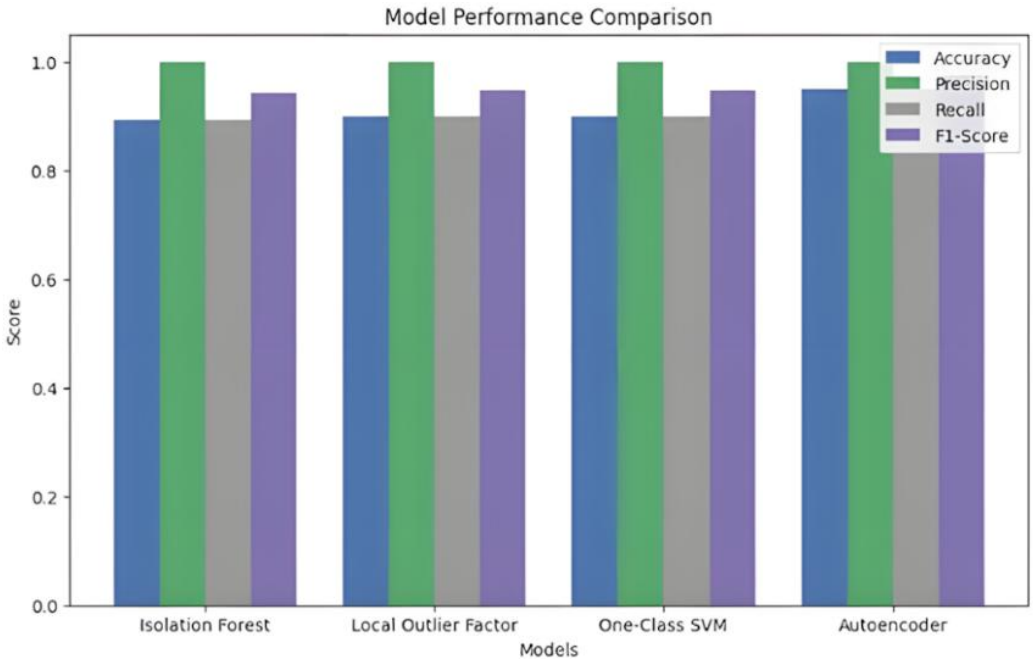


Figure 3: Comparison of different metrics with 4 ML models.

The Autoencoder achieved superior performance compared to all other models while modelling civic energy patterns through nonlinear approaches. The ensemble system shows excellent precision but reduced accuracy levels, which indicates its capability to minimize false alarms in detection systems during critical energy utilization.

4.2 Anomaly Detection Visualization

Visualization offered a qualitative understanding of how different models behave. In Figure 4, a scatter plot compares anomaly detection (0) and normal (1) predictions of Isolation Forest, Local Outlier Factor (LOF), One Class SVM, and Autoencoder models with offsets for clarity. It shows distinct anomaly clusters for the Autoencoder. Anomaly Detection Plots implemented scatter plots between sample indices and feature values, which were modified through the application of various models, to visualize anomaly distributions. Autoencoder detected the most consistent anomalous points (red points) in the data, while Isolation Forest and LOF produced more scattered detection results due to their global and local detection strategies, respectively.

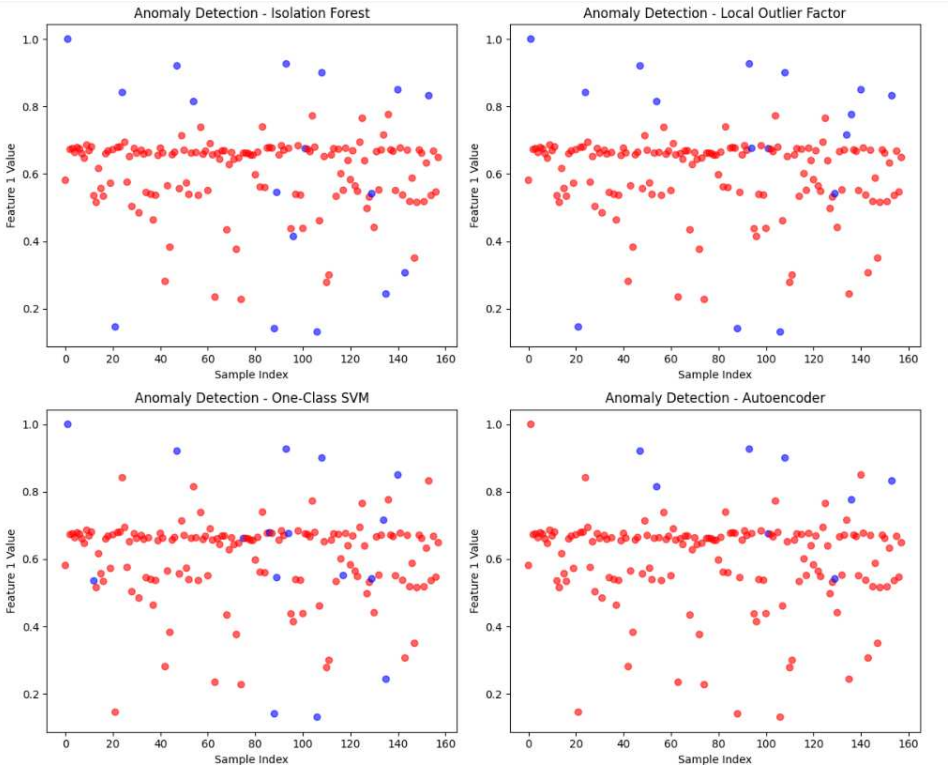


Figure 4: Scatter plot comparing anomaly detection with different ML models.

Figure 5 shows TSNE 2D scatter plot visualizing anomaly clusters in X_test, with colors representing predictions (red for anomalies, blue/purple for normal), highlighting the Autoencoder's clustering. X_test TSNE projection showed abnormal data points in red clusters separate from regular data points in blue/purple areas. Tight groupings of anomalies emerged through the Autoencoder and One-Class SVM, which indicated superior separation when compared to the dispersed clusters found in Isolation Forest and LOF. Figure 6 shows that the Autoencoder exhibits an MSE score range between 0 and 2, which clearly demonstrates a threshold compared to the more restricted score ranges of alternative models. It shows box plots of the distribution of anomaly scores generated by four unsupervised anomaly detection models: isolation forest, local outlier factor (LOF), one-class SVM, and autoencoder. Each box plot depicts the central tendency, distribution, and probable outliers of the anomaly scores calculated by each technique on the test dataset.

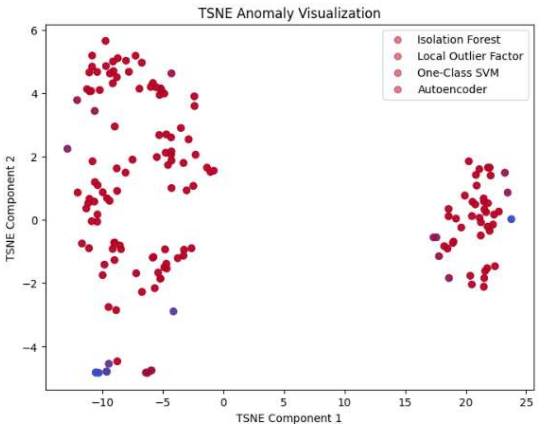


Figure 5: TSNE 2D scatter plot visualizing anomaly clusters.

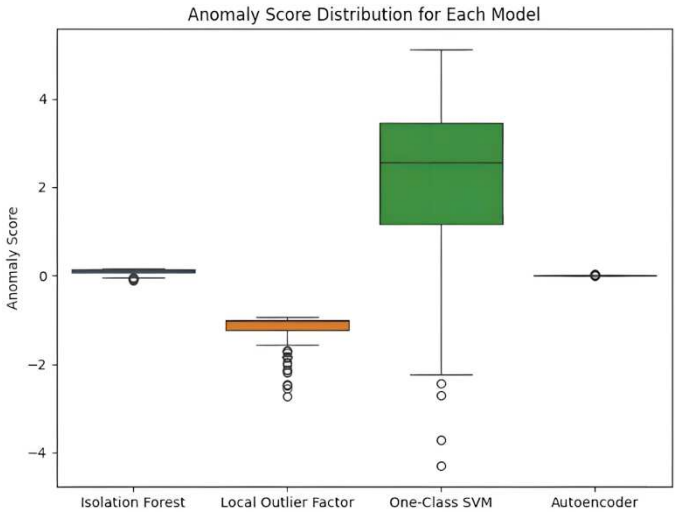


Figure 6: Box plots of anomaly scores.

4.3.1 General Structure and Interpretation of Box Plots

A box plot illustrates the statistical distribution of data through five key components: the median, represented as a central line within the box; the first quartile (Q1) and third quartile (Q3), which define the box boundaries enclosing the middle 50% of the dataset; and the whiskers, which extend from Q1 and Q3 up to 1.5 times the interquartile range (IQR) to represent the primary spread of the data. Points located outside the whiskers are identified as outliers, indicating values that deviate significantly from normal distribution patterns.

Color in Figure 6 helps to understand how different models behave. The One-Class SVM plot (marked in green) occupies the widest range due to an enlarged IQR and considerable score variability, which explains why it has the largest range in the plot. On the contrary, the box plots of Isolation Forest, Local Outlier Factor (LOF), and Autoencoder appear visually similar and share a standard tone and thus are statistically similar, which reinforces comparability. Black markers denote outlier points across all models, which indicate exceptional anomaly scores.

4.3.2 Comparative Analysis of Model Box Plots

Isolation Forest: The Isolation Forest plot shows a narrow box centered near zero. This indicates little variation and a stable distribution of anomalous scores. There is singular upper outlier, which suggests one highly anomalous data instance. The symmetry of the box confirms uniformity in the tree-based configuration and consistent, reliable anomaly detection.

Local Outlier Factor (LOF): The LOF plot indicates consistency in the detection of local outliers: a narrow box with its median slightly under zero, which suggests a negative scoring convention, and hence, a score lower than zero based on local density variations, is consistent. The compact IQR confirms that LOF detects high precision with localized dense data region disproportionate deviations.

One-Class SVM: The One-Class SVM shows the widest box, with the whiskers extending from around -2 to 4. This indicates variability and sensitivity to several levels of anomalous detection. The numerous outliers extending beyond the score range suggest the model is able to capture varied irregular patterns, albeit with an increased risk of false positives. The wide IQR corresponds with the model's kernel-based mechanism of hyperplane separation of normal and abnormal samples.

Autoencoder: From the provided Autoencoder plot, it can be seen a dense and tightly bounded box with the median value close to zero and MSE ranging from 0 to 2. Because there are short whiskers with a single upper outlier, it can be said there is strong consistency of reconstruction and effective anomaly differentiation. As the Autoencoder correctly reconstructs normal data, the narrow IQR shows the system identifies rare energy inefficiencies reconstruction with high peaks of error as inefficiencies, high reconstruction errors.

4.4 Interpretability Analysis

In Figure 7, a SHAP analysis on the Autoencoder model showed the most important features in the dataset, as the Autoencoder captures important features. Feature 2 and Feature 6, with mean (|SHAP|) values of 0.25 and 0.20, highlight these features as the most important for anomaly detection, as they reference peak usage times and high energy consumption patterns within the civic energy dataset, integrating the most energy consumption patterns, as these patterns trigger anomaly detection.

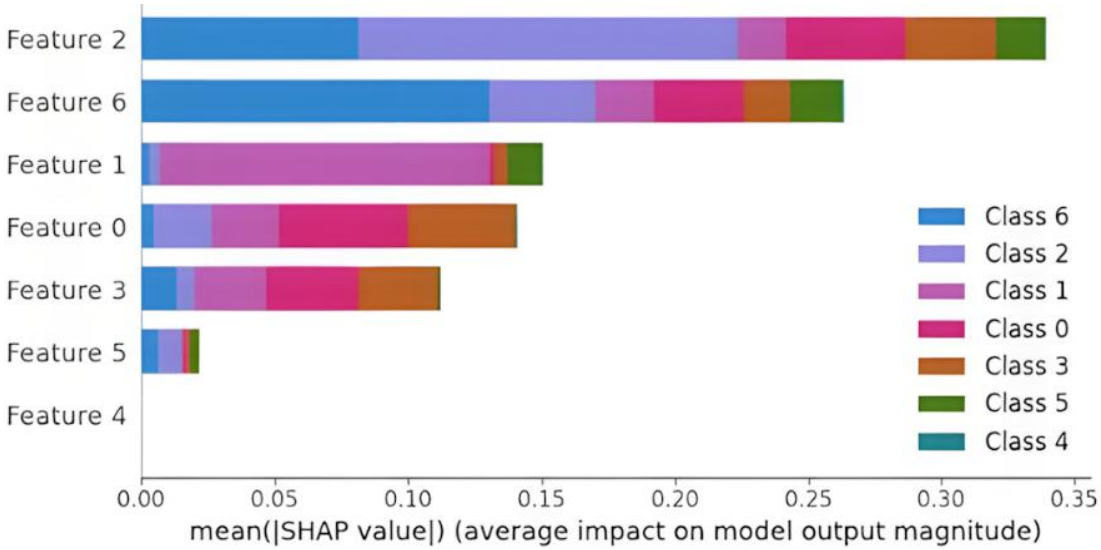


Figure 7: SHAP summary plot showing feature impacts on Autoencoder anomaly detection, with Feature 2 and Feature 6 as the most influential.

4.5 Discussions

The Autoencoder anomaly detection methods achieve an F1-Score of 0.9740, demonstrating a successful handling of inefficient usage pattern civic energy anomalies. The detection ensemble approach shows a balanced strength of Isolation Forest and LOF, yet its accuracy (0.8861) suggests possibly conservative voting behaviour. The Autoencoder identifies urban energy-conserving peak demand anomalies given its effective separation of anomalies using the 95th percentile threshold, which validated anomaly detection using SHAP analysis on Feature 2 and Feature 6. The synthetic ground truth data produces a total precision measurement (1.0000) due to data imbalance and the lack of time-series analysis. Inspectors are inhibited from identifying temporal patterns in the anomalies. Yet the ensemble method from our research provides strong reliability, and the SHAP interpretability framework provides actionable insights to public administrators. Incorporation of genuine labelled data and dynamic thresholding will enhance the effectiveness of civic energy issues in real time.

5. Conclusions

An ensemble-based anomaly detection framework integrating Isolation Forest, Local Outlier Factor, One-Class SVM, and Autoencoder techniques is developed in this study to identify patterns of irregularities in urban power consumption. Intelligent feature selection and careful data preprocessing improved system reliability, while SHAP analysis achieved interpretability by demonstrating feature contributions and impacts. Among the anomaly detection models, the Autoencoder outperformed others, achieving the highest scores of 0.9494 and 0.9740 on accuracy and F1 metrics, respectively, and surpassing the performance of conventional techniques. Although a slightly lower accuracy of 0.8861 was observed, the ensemble model achieved perfect precision of 1.0000, signaling almost no false positives. SHAP analysis results demonstrate that Features 2 and 6, which capture most of the peak energy consumption, greatly contributed to accurate anomaly detection. This framework identifies wasteful consumption in urban power systems, thereby assisting in the sustainable power management of civic infrastructure and the attainment of UN SDG 7 (Affordable and Clean Energy). In future work, real-world time-series datasets will be used to enable real-time analysis and monitor adaptive time-series for the problem.

Declarations

Ethical Approval and Consent to Participate

This study did not involve human participants, animals, or any personal data. Therefore, ethical approval and informed consent were not required.

Consent for Publication

All authors have reviewed the manuscript and have given their consent for publication.

Availability of Data and Materials

The datasets generated and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

1. R. Daniel Sorin, “Enhancing Urban Sustainability: The Role of Artificial Neural Networks in Energy Efficiency and Smart City Development,” International Journal of Science and Research (IJSR), vol. 12, no. 12, pp. 1620–1624, Dec. 2023, doi: 10.21275/sr231221174419.
2. Y. WANG, Y. LIU, and C. XU, “Application of outlier customer meaning analysis in quality management,” Journal of Computer Applications, vol. 29, no. 11, pp. 3077–3979, Dec. 2009, doi: 10.3724/sp.j.1087.2009.03077.
3. V. Chandola, A. Banerjee, and V. Kumar, “Anomaly detection,” ACM Computing Surveys, vol. 41, no. 3, pp. 1–58, Jul. 2009, doi: 10.1145/1541880.1541882.

4. P. Olasolo-Alonso, L. M. López-Ochoa, J. Las-Heras-Casas, and L. M. López-González, “Energy Performance of Buildings Directive implementation in Southern European countries: A review,” Energy and Buildings, vol. 281, p. 112751, Feb. 2023, doi: 10.1016/j.enbuild.2022.112751.
5. R. Moghaddass and J. Wang, “A Hierarchical Framework for Smart Grid Anomaly Detection Using Large-Scale Smart Meter Data,” IEEE Transactions on Smart Grid, vol. 9, no. 6, pp. 5820–5830, Nov. 2018, doi: 10.1109/tsg.2017.2697440.
6. F. Xu and R. Moghaddass, “A scalable Bayesian framework for large-scale sensor-driven network anomaly detection,” IISE Transactions, vol. 55, no. 5, pp. 445–462, Apr. 2022, doi: 10.1080/24725854.2022.2037792.
7. S. Ghandehariun, A. M. Ghandehariun, and N. B. Ziabari, “Performance prediction and optimization of a hybrid renewable-energy-based multigeneration system using machine learning,” Energy, vol. 282, p. 128908, Nov. 2023, doi: 10.1016/j.energy.2023.128908.
8. A. Majnoon and A. Saifoddin Asl, “AI-Driven Energy Optimization Enhancing Efficiency in Urban Environments with Hybrid Machine Learning Models,” 2024, doi: 10.2139/ssrn.4960870.
9. F. T. Liu, K. M. Ting, and Z.-H. Zhou, “Isolation Forest,” 2008 Eighth IEEE International Conference on Data Mining, pp. 413–422, Dec. 2008, doi: 10.1109/icdm.2008.17.
10. M. M. Breunig, H.-P. Kriegel, R. T. Ng, and J. Sander, “LOF,” ACM SIGMOD Record, vol. 29, no. 2, pp. 93–104, May 2000, doi: 10.1145/335191.335388.
11. B. Schölkopf, J. C. Platt, J. Shawe-Taylor, A. J. Smola, and R. C. Williamson, “Estimating the Support of a High-Dimensional Distribution,” Neural Computation, vol. 13, no. 7, pp. 1443–1471, Jul. 2001, doi: 10.1162/089976601750264965.
12. J. Heaton, “Ian Goodfellow, Yoshua Bengio, and Aaron Courville: Deep learning,” Genetic Programming and Evolvable Machines, vol. 19, no. 1–2, pp. 305–307, Oct. 2017, doi: 10.1007/s10710-017-9314-z.
13. Z.-H. Zhou, “Ensemble Methods,” Jan. 2025, doi: 10.1201/9781003587774.
14. Z. Zhang, Y. Chen, H. Wang, Q. Fu, J. Chen, and Y. Lu, “Anomaly detection method for building energy consumption in multivariate time series based on graph attention mechanism,” PLOS ONE, vol. 18, no. 6, p. e0286770, Jun. 2023, doi: 10.1371/journal.pone.0286770.
15. S. K. Singh, Y.-S. Jeong, and J. H. Park, “A deep learning-based IoT-oriented infrastructure for secure smart City,” Sustainable Cities and Society, vol. 60, p. 102252, Sep. 2020, doi: 10.1016/j.scs.2020.102252.
16. M. Yang and D. Sharma, “The Spatiality and Temporality of Electricity Reform: A Comparative and Critical Institutional Perspective,” Energy Research & Social Science, vol. 60, p. 101327, Feb. 2020, doi: 10.1016/j.erss.2019.101327.
17. B. Ng, “Survey of Anomaly Detection Methods,” Office of Scientific and Technical Information (OSTI), Oct. 2006, doi: 10.2172/900157.
18. D. Borgetto, G. Da Costa, J.-M. Pierson, and A. Sayah, “Energy-aware resource allocation,” 2009 10th IEEE/ACM International Conference on Grid Computing, pp. 183–188, Oct. 2009, doi: 10.1109/grid.2009.5353063.

19. C. Xu and H. Chen, “A hybrid data mining approach for anomaly detection and evaluation in residential buildings energy data,” Energy and Buildings, vol. 215, p. 109864, May 2020, doi: 10.1016/j.enbuild.2020.109864.
20. P. Kurtser, K. Feng, T. Olofsson, and A. De Andres, “One-class anomaly detection through color-to-thermal AI for building envelope inspection,” Energy and Buildings, vol. 328, p. 115052, Feb. 2025, doi: 10.1016/j.enbuild.2024.115052.
21. T. Liu, Z. Zhou, and L. Yang, “Layered isolation forest: A multi-level subspace algorithm for improving isolation forest,” Neurocomputing, vol. 581, p. 127525, May 2024, doi: 10.1016/j.neucom.2024.127525.
22. M. Ibrahim, A. Alsheikh, F. Awaysheh, and M. Alshehri, “Machine Learning Schemes for Anomaly Detection in Solar Power Plants,” Energies, vol. 15, no. 3, p. 1082, Feb. 2022, doi: 10.3390/en15031082.
23. A. Chen, J. Wu, and H. Zhang, “FIAD: Graph anomaly detection framework based feature injection,” Expert Systems with Applications, vol. 259, p. 125216, Jan. 2025, doi: 10.1016/j.eswa.2024.125216.
24. Z. Chen, F. Xiao, F. Guo, and J. Yan, “Interpretable machine learning for building energy management: A state-of-the-art review,” Advances in Applied Energy, vol. 9, p. 100123, Feb. 2023, doi: 10.1016/j.adapen.2023.100123.
25. H. Luo, Y. Zheng, K. Chen, and S. Zhao, “Probabilistic Temporal Fusion Transformers for Large-Scale KPI Anomaly Detection,” IEEE Access, vol. 12, pp. 9123–9137, 2024, doi: 10.1109/access.2024.3353201.
26. J. Zhang, X. Jin, Y. Liang, and J. You, “Anomaly Detection Method for Smart Distribution Grids Using Autoencoders,” 2024 9th International Conference on Intelligent Computing and Signal Processing (ICSP), pp. 398–402, Apr. 2024, doi: 10.1109/icsp62122.2024.10743913.
27. J. Duan, “Deep learning anomaly detection in AI-powered intelligent power distribution systems,” Frontiers in Energy Research, vol. 12, Mar. 2024, doi: 10.3389/fenrg.2024.1364456.
28. U. A. Usmani, I. Abdul Aziz, J. Jaafar and J. Watada, "Deep Learning for Anomaly Detection in Time-Series Data: An Analysis of Techniques, Review of Applications, and Guidelines for Future Research," in *IEEE Access*, vol. 12, pp. 174564-174590, 2024, doi: 10.1109/ACCESS.2024.3495819.
29. L. Mascali et al., “A machine learning-based Anomaly Detection Framework for building electricity consumption data,” Sustainable Energy, Grids and Networks, vol. 36, p. 101194, Dec. 2023, doi: 10.1016/j.segan.2023.101194.
30. W. Zhang, X. Li, H. Ma, Z. Luo, and X. Li, “Universal Domain Adaptation in Fault Diagnostics With Hybrid Weighted Deep Adversarial Learning,” IEEE Transactions on Industrial Informatics, vol. 17, no. 12, pp. 7957–7967, Dec. 2021, doi: 10.1109/tii.2021.3064377.
31. L. Göke and M. Kendziorski, “Adequacy of time-series reduction for renewable energy systems,” Energy, vol. 238, p. 121701, Jan. 2022, doi: 10.1016/j.energy.2021.121701.
32. Z. Zhang, Y. Chen, H. Wang, Q. Fu, J. Chen, and Y. Lu, “Anomaly detection method for building energy consumption in multivariate time series based on graph attention

mechanism," PLOS ONE, vol. 18, no. 6, p. e0286770, Jun. 2023, doi: 10.1371/journal.pone.0286770.
33. S. K. Singh, Y.-S. Jeong, and J. H. Park, "A deep learning-based IoT-oriented infrastructure for secure smart City," Sustainable Cities and Society, vol. 60, p. 102252, Sep. 2020, doi: 10.1016/j.scs.2020.102252.