

A Multitask Framework for Optimizing Smart Grid Energy Consumption Using RegClassXNet and Dynamic Cluster Adjustment

Hegang Zong¹, Bihui Chen¹, Dingli Li¹, Chengfu Liu¹, Yang Cai¹

¹Huaneng Lancang River Hydropower Inc, Kunming City, Yunnan Province, 650214, China

E-mail: 18788548981@163.com

Keywords: Electricity consumption, energy management, sustainable cities, smart grids, deep learning, data-driven optimization

Received: December 20, 2024

With urbanization accelerating electricity demand, advanced energy management is vital for building sustainable cities. This paper presents a novel framework using real-world data from commercial, residential, and industrial buildings over six years to optimize electricity consumption. The proposed RegClassXNet model, integrating EfficientNet, Xception, and Swin-Transformer, performs multitask predictions for both classification and regression objectives. Proportional Dynamic Cluster Adjustment (PDCA) is introduced to address data imbalance, and a hybrid attribute refinement process synthesizes relevant features to enhance predictive accuracy. Our model achieves 95.0% R-squared, 2.1 MAE, 1.8 RMSE, and 3.2 MSE, significantly outperforming existing methods such as CNN, LSTM, and RF. New stability metrics, including Label Variability Consistency Index (LVCI), Temporal Prediction Stability Measure (TPSM), and Output Correlation Coefficient (OCC), ensure robust and consistent predictions. The framework was evaluated on a dataset comprising 500,000 energy consumption records, utilizing a distributed training approach on a high-performance GPU cluster. Simulations illustrate the framework's capability to optimize energy usage across building types, adjust for environmental impacts, and support effective energy-saving strategies. This work offers a transformative approach to sustainable energy management, paving the way for adaptive, data-driven smart grid systems.

Povzetek: Predstavljen je večopravilen okvir RegClassXNet za optimizacijo porabe energije v pametnih omrežjih, ki s kombinacijo EfficientNet, Xception in Swin-Transformerja dosega visoko točnost in stabilnost napovedi.

1 Introduction

The rise in global energy demand, driven by urbanization and population expansion, poses challenges to conventional power grid systems. Smart Grids (SG) enhance energy management by providing more reliability, efficiency, and sustainability via integrating IoT, AI, and big data analytics for real-time monitoring, demand response, and load forecasting. The intricacy of energy demand and urban density hampers energy management. Advanced energy management systems optimize power consumption and operational efficiency, while predictive analytics improve energy utilization in smart buildings for lighting, HVAC, and other requirements [2]. The fluctuation of renewable sources such as solar and wind exacerbates energy management challenges, necessitating effective load forecasting and storage solutions [3, 4].

Intelligent energy management employs IoT devices, including smart meters and sensors, to monitor and optimize energy use. These devices enhance load forecasting and demand-side management by gathering comprehensive consumption data [5]. Artificial intelligence methodologies, including deep learning (DL) and machine learning

(ML), are used to predict energy usage, discern trends, and recommend energy conservation measures [6, 7]. However, managing extensive datasets and mitigating overfitting persist in significant problems, underscoring the need for efficient feature selection and model optimization [8]. In smart cities, energy management encompasses residential, commercial, and industrial sectors, using energy-efficient technology and renewable resources [9, 10, 11]. Smart grids and microgrids improve resilience and enable local energy trade. Load forecasting techniques, such as LSTM networks, enhance prediction accuracy across various time frames, adapting to swiftly evolving consumption patterns [12].

AI-driven energy management may optimize the supply-demand equilibrium via demand response tactics, hence decreasing peak consumption [13, 14, 15]. Notwithstanding progress, scalable and resilient solutions are crucial for managing the fluctuation of renewable energy sources and safeguarding data security and privacy in IoT systems [16]. Hybrid AI models, real-time algorithm optimization, and edge computing are being investigated to address these difficulties. This article presents the following contributions.

1. Innovative Multilabel Framework for Energy Man-

agement: Created a novel methodology that integrates power consumption predictions, energy savings potential assessment, and other classification variables into a multilabel classification and regression framework. This comprehensive methodology elucidates the interconnections among many energy measures, yielding enhanced insights for optimization and sustainability.

2. **Innovative RegClassXNet Model for Multilabel Prediction:** Presented the RegClassXNet architecture, which amalgamates EfficientNet, Xception, and Swin-Transformer layers to manage both categorical and continuous outputs. The model exhibits exceptional efficacy in forecasting various energy-related results, enhancing energy management tactics across several building types and contexts.
3. **Advanced Techniques for Data Balancing and Feature Engineering:** Introduced the Proportional Dynamic Cluster Adjustment (PDCA) to tackle data imbalance and a Hybrid Attribute Refinement and Synthesis method for creating relevant features. These strategies augment data quality and bolster the model's predictive capability.
4. **Implementation of Novel Assessment Metrics:** Introduced three innovative metrics—Label Variability Consistency Index (LVCI), Temporal Prediction Stability Measure (TPSM), and Output Correlation Coefficient (OCC)—to enhance the evaluation of multilabel classification and regression tasks, focussing on performance consistency, prediction stability, and output interrelations.
5. **This study integrates energy consumption forecasting with carbon emission reduction analysis, providing practical information for sustainability initiatives. It delineates essential characteristics and timeframes for focused actions, facilitating demand response techniques and activities to reduce environmental impact.**

The remainder of the work follows: Section 2 presents the literature overview and a summary of pertinent energy management and predictive modeling research. Section 3 describes the suggested approach, the assessment criteria, the RegClassXNet model, and data preparation. Section 4 addresses the simulation results and analysis, contrasting the proposed strategy's performance with current techniques. The last part emphasizes results and possible paths for further investigations.

2 Related work

Research on optimizing electricity consumption for power use is crucial for intelligent energy management systems. Many research studies have used machine learning (ML) and deep learning (DL) to improve accuracy and efficiency in predicting.

Author in [17, 18] presents a hybrid short-term load forecasting technique combining FFNN, RNN, LSTM, and SVR. The hybrid model outperformed separate models in trials to increase electrical system prediction accuracy. The hybrid system's complexity may raise computing costs and training time, rendering it unsuitable for real-time applications. [19] thoroughly reviews machine learning techniques for load demand prediction. The goal was to find the best load forecasting algorithms using LSTM, SVR, and Random Forest. The research showed that machine learning can forecast load demand, although data quality and pre-processing affect accuracy. This vulnerability emphasizes the need for proper data management in load forecasting research. Another notable contribution is [20], which develops a DLSTM model for short-term load forecasting [21]. Weather-related characteristics are included in the model to enhance Panama's energy demand forecast [22]. The paper indicates that the DLSTM improves accuracy but requires significant training data, which may not be available in all locations. The work in [23] explores medium-term regional electrical load forecasting utilizing RF and LSTM ensemble learning approaches. A 12-year dataset is used to assess model performance in this research. With a MAPE of 6.46%, the LSTM model outperformed conventional approaches. However, the model's heavy use of historical data raises questions about its applicability in data-poor places.

Researchers in [24] introduced a novel multi-sequence LSTM-RNN model optimized using metaheuristic algorithms like GA and PSO. This research selected suitable hyperparameters to improve the LSTM model for electric load forecasting. A multi-sequence LSTM model from metaheuristics outperformed benchmark models and other machine learning methods like SVR and ANN [25]. However, hyperparameter adjustment remains difficult, which may restrict the model's real-world applicability. Research by [26] highlights the effectiveness of CNN and LSTM deep learning approaches for short-term load forecasting [27]. Automation of load data breakdown was sought to increase forecasting accuracy. While deep learning algorithms may capture complicated energy use patterns, massive datasets for training are a significant obstacle.

Using their strengths, the researcher in [28, 29] developed a hybrid model integrating LSTM and CNN for short-term load forecasting. The hybrid model outperformed standard approaches in the research, but its complexity may raise training durations and computer resource demands, making implementation difficult. The paper in [30, 31] explores AI-powered energy optimization in intelligent cities, focussing on a framework integrating machine learning and deep learning models for real-time load forecasting. The recommended approaches increased energy efficiency significantly [32]. Integrating models and data sources complicates system management and operational efficiency [?]. Another research used machine learning for intelligent grid short-term demand predictions. The goal was to assess the accuracy of these approaches in predicting electricity con-

sumption, considering environmental factors like Temperature and humidity [34]. Machine learning enhanced predicting accuracy, but quality input data and feature selection remained challenges. The study in [35] created a hybrid load forecasting model combining LSTM and SVR. Combining the qualities of both methods improved forecast accuracy. The findings showed enhanced predicting accuracy. However, hybridization complexity makes implementation difficult, especially regarding computing efficiency. Energy management solutions in intelligent manufacturing use IoT technology to optimize efficiency. The Energy Management System (EMS) involves reviewing existing practices and creating concrete strategies to increase energy efficiency [36]. The potential for energy savings in production is limited by data integration and real-time analysis issues [37].

In intelligent buildings and industries, AI-driven forecasting and energy management systems are essential to optimizing power use. While various state-of-the-art models have been explored for electricity load forecasting, limitations such as imbalanced data handling, computational overhead, and lack of generalization persist. Table 1 provides a comparative summary of these existing approaches, highlighting their key outcomes and limitations, thus justifying the need for the proposed RegClassXNet framework. Despite significant progress, intelligent energy management frameworks must overcome feature redundancy, computational complexity, and real-time adaptation. This project seeks to establish a comprehensive framework using advanced AI approaches to optimize power demand forecasting and improve innovative energy system sustainability using lessons from previous investigations.

3 Proposed method

The proposed method model uses IoT-based data gathering, advanced machine learning, and multilabel classification regression to optimize power consumption forecasts and energy management. The model uses a 2018–2024 dataset of hourly energy use and environmental indicators from residential, commercial, and industrial buildings in Southern California. The RegClassXNet model combines EfficientNet, Xception, and Swin-Transformer layers with data preprocessing, feature selection, and hybrid modeling. This hybrid architecture captures complex data linkages and interactions to estimate energy-related objectives, including energy consumption, energy savings potential, peak demand reduction, and carbon emission categories.

3.1 Data collection and description

The dataset used in this study was collected from a comprehensive energy monitoring system implemented across multiple commercial, residential, and industrial buildings in Southern California from January 2018 to January 2024 [38]. This region was selected due to its diverse climate conditions and varying energy usage patterns, which offer a

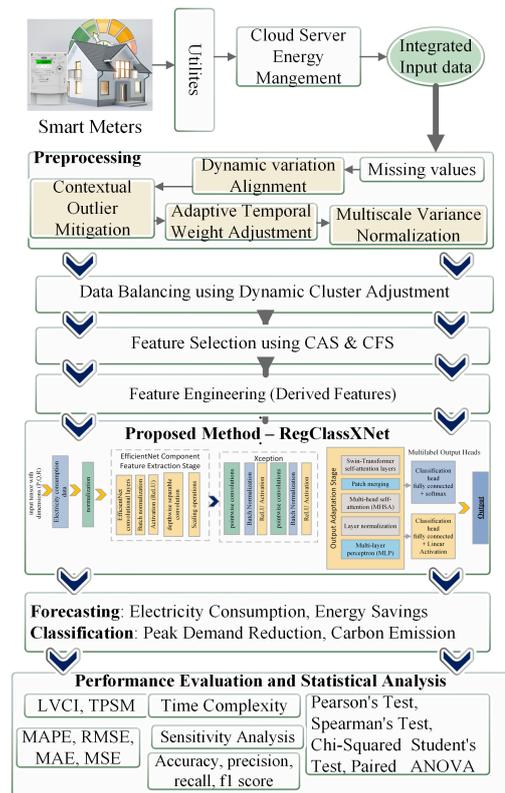


Figure 1: Proposed framework

rich context for analyzing electricity consumption and optimization opportunities. The data were sourced from smart meters, IoT sensors, and energy management systems installed in over 100 facilities, capturing detailed hourly energy usage and environmental metrics. The collection process involved integrating data from regional utility companies, building management systems, and independent energy monitoring providers. The dataset encompasses different seasons and significant events, such as public holidays and extreme weather conditions, to reflect real-world variations in energy demand. This comprehensive dataset is valuable for exploring electricity consumption trends, optimizing energy usage, and assessing the potential for energy savings through artificial intelligence-driven approaches. Including diverse building types and conditions ensures the findings broadly apply across various regional settings.

3.2 Data preprocessing

Due to the dynamic nature of Southern California power usage, the dataset is preprocessed to enhance quality and consistency. These unique approaches are introduced: Dynamic variation Alignment (DVA) aligns variation across periods to handle power usage and other environmental factor swings. This adjusts feature variability depending on sliding window mean values. Let y_j represent a feature value at time j , and θ_{m_j} and η_{m_j} represent the mean and standard deviation of the sliding window m of size M . Calculating adjusted y'_j :

Table 1: Comparison of state-of-the-art studies on electricity load forecasting

Ref	Method Used	Objective Achieved	Limitations
[17]	Hybrid FFNN, RNN, LSTM, SVR	Enhance prediction accuracy for electrical systems	Increased computational costs and longer training times
[19]	LSTM, SVR, RF	Identify effective algorithms for load forecasting	Sensitivity to data quality and pre-processing techniques
[21]	DLSTM	Improve prediction accuracy incorporating weather features	Challenges due to the need for extensive training data
[23]	Ensemble RF, LSTM	Evaluate model performance over a 12-year dataset	Reliance on extensive historical data limits applicability
[24]	Multi-sequence LSTM-RNN optimized by GA, PSO	Enhance LSTM model performance for load forecasting	Complexity of hyperparameter tuning limits practicality
[26]	CNN, LSTM	Automate load data decomposition to improve accuracy	Requirement for large datasets for effective training
[29]	Hybrid LSTM, CNN	Leverage strengths of both architectures for forecasting	Increased complexity can lead to longer training times
[31]	AI methods integrating ML and DL	Develop a framework for real-time load forecasting	Complex integration of various models complicates management
[34]	Various ML techniques	Effectively predict electricity demand incorporating external variables	Dependence on quality input data and feature selection
[35]	Hybrid LSTM, SVR	Enhance prediction accuracy using hybrid model strengths	Complexity of hybridization poses implementation challenges
[36]	EMS in IoT systems	Optimize energy efficiency through systematic evaluation	Ongoing challenges in data integration and real-time analysis

Table 2: Dataset features overview

S.No	Features	Short Description
1	Timestamp	Date and time of the recorded data, sampled hourly from 2018 to 2024.
2	Building Type	Type of building, such as residential, commercial, or industrial.
3	Energy Consumption (kWh)	Total electricity consumption recorded in kilowatt-hours.
4	Temperature (°C)	Outdoor Temperature recorded during consumption.
5	Humidity (%)	Relative humidity levels during energy consumption.
6	Occupancy Rate (%)	Estimated percentage of occupied spaces in the building.
7	Lighting Consumption (kWh)	Electricity used for lighting within the building.
8	HVAC Consumption (kWh)	Energy used by heating, ventilation, and air conditioning systems.
9	Energy Price (\$/kWh)	Cost of electricity at the time of consumption.
10	Carbon Emission Rate (g CO2/kWh)	Associated carbon emissions from energy consumption.
...
47	Thermal Comfort Index	Index indicating the perceived comfort level based on Temperature and humidity.
48	Energy Savings Potential (%)	Predicted potential savings in energy consumption.
49	Peak Demand Reduction Indicator	Binary indicator of whether peak demand is reduced.
50	Carbon Emission Reduction Category	Categorical value indicating levels of reduction in carbon emissions.

$$y'_j = \frac{y_j - \theta_{m_j}}{\eta_{m_j} + \delta} \cdot \sqrt{M} \quad (1)$$

where a modest constant δ is introduced to avoid zero division. This modification improves consistency by aligning each data segment with the overall temporal variance. By analyzing the links between outliers in the dataset across different periods and building types, Contextual Outlier Mitigation (COM) can identify and correct them. To reduce the effect of superfluous outliers, this method takes feature correlations into account rather than depending just on statistical outlier identification. The corrected value y'_j is calculated as follows:

$$y'_j = \begin{cases} y_j, & \text{if } |y_j - \bar{y}_q| \leq \beta \cdot \gamma_q \\ \bar{y}_q + \zeta \cdot (y_j - \bar{y}_q), & \text{otherwise} \end{cases} \quad (2)$$

the contextual mean and standard deviation for a specific feature category q are represented by \bar{y}_q and γ_q , respectively, and scaling factors β and ζ are used to modify deviations. Adaptive Temporal Weight Adjustment (ATWA) uses time-of-day weights to adjust for energy consumption fluctuations caused by seasonal influences and exceptional events. This approach highlights crucial consumption periods by shifting the weights to highlight peak demand and off-peak consumption periods. Here is the definition of the weight u_j for time point j :

$$u_j = 1 + \rho \cdot \sin\left(\frac{2\pi\omega_j}{\Omega}\right) \quad (3)$$

The time index is ω_j , the daily hours are Ω , and the scaling factor is ρ , which modifies the weight range. This adjustment covers data on daily and seasonal cycles. Normalizing Using Multiscale Variance Normalisation (MVN), normalizing data across scales corrects feature value imbalances. Smoothing variances with multi-resolution improves feature value stability. Normalize each feature's y'_j :

$$y'_j = \frac{y_j - \mu_n}{\sqrt{\sigma_n^2 + \lambda \cdot \sigma_{\text{local}}^2}} \quad (4)$$

where μ_n and σ_n represent the larger-scale mean and standard deviation. For n , σ_{local} reflects local variance inside a narrower window and λ balances global and local normalization. Harmonic Pattern Enhancement (HPE) stabilizes

data with cyclical patterns like energy usage from everyday activities using harmonic functions. The increased value y'_j is:

$$y'_j = y_j + \tau \cdot \sum_{l=1}^L b_l \cos\left(\frac{2\pi l \omega_j}{\Omega} + \psi_l\right) \quad (5)$$

where L is the number of harmonic components, b_l represents the amplitude, ψ_l is the phase shift, Ω is the period, and τ is the smoothing factor. These preprocessing steps improve the dataset by addressing outliers, normalizing variations across scales, adjusting for temporal effects, and enhancing cyclical patterns, ensuring the data is prepared for effective modeling and optimization of electricity consumption through AI.

3.3 Conversion of imbalanced data to balanced data using dynamic cluster adjustment (PDCA)

This study's dataset is imbalanced across characteristics and target labels, which may bias model training and performance. Proportional Dynamic Cluster Adjustment (PDCA) is a new data balancing mechanism [41]. This approach dynamically modifies sample weights depending on cluster data point density to appropriately represent minority groups without oversampling. Clustering, proportional correction, and dynamic resampling comprise the method. The dataset is clustered using density-based clustering. Let P_m represent the sample count in each cluster D_m , and P_{sum} represent the overall sample count in the dataset. Cluster proportional weight is computed as:

$$\omega_m = \frac{1}{1 + \exp\left(-\gamma \cdot \left(\frac{P_{\text{max}} - P_m}{P_{\text{sum}}}\right)\right)} \quad (6)$$

A scaling factor γ governs the pace of adjustment, whereas P_{max} represents the greatest cluster size. This function weights smaller clusters higher, boosting their resampling representation. After computing weights, dynamic resampling generates synthetic samples for low-density clusters and downsamples high-density clusters. For each cluster, Q_m synthetic samples are generated:

$$Q_m = \left(\frac{\omega_m \cdot P_{\text{sum}}}{\sum_{n=1}^L \omega_n}\right) - P_m \quad (7)$$

where L is the cluster count. Linear interpolation between data points in each cluster and a random perturbation factor generates synthetic samples with variability. Combining the original data points and the synthetic samples with an adjustment factor to avoid cluster over-representation yields a balanced dataset. The correction factor λ_m for cluster D_m is calculated as:

$$\lambda_m = \frac{P_m + Q_m}{P_{\text{sum}} + \sum_{n=1}^L Q_n} \quad (8)$$

This approach uniformizes data distribution, minimizing dataset imbalance while keeping data properties. The research balances the dataset via Proportional Dynamic Cluster Adjustment, improving the model's generalization across classes.

3.4 Hybrid attribute refinement and synthesis

Hybrid Attribute Refinement and Synthesis improves modeling feature quality and relevance [42]. This method uses repeated selection and sophisticated synthesis to find the most important traits and create new, valuable features from current ones. The Composite Adaptive Selector (CAS) selects attributes, and the Contextual Feature Synthesis (CFS) generates new features. The Composite Adaptive Selector (CAS) optimizes feature selection using statistical relevance, density-based importance, and evolutionary divergence criteria. We create an aggregated significance score for each attribute A_p :

$$\Gamma_p = \delta \cdot \Theta_p + \kappa \cdot \Pi_p + \sigma \cdot \Lambda_p \quad (9)$$

The statistical relevance of a feature is calculated using mutual information, while the density-based significance score is derived from local density variations within clusters. The evolutionary divergence score is calculated based on feature distribution shifts over time. The weighting factors δ , κ , and σ regulate the contribution of each component. Select features with the highest aggregated score Γ_p for refining. The innovative Contextual Feature Synthesis (CFS) method improves chosen characteristics. This technique transforms and combines contextually related information to create new attributes. New feature B_q is synthesized using:

$$B_q = \sum_{r=1}^U \phi_r \cdot z_r + \frac{1}{\nu_q + \zeta} \cdot \sum_{s=1}^V (\omega_s \cdot x_s) \quad (10)$$

z_r as an original feature, ϕ_r as a transformation coefficient, ν_q as a context factor for feature B_q , and ζ as a small constant to prevent zero division. The second term uses x_s features and ω_s scaling coefficients to capture attribute interactions. The equation generates new characteristics that represent complicated data connections. DAS improves the interpretability and consistency of synthesized features by adjusting the scale of freshly created characteristics. Calculate the scaled value \hat{B}_q of the synthesised feature B_q :

$$\hat{B}_q = \frac{B_q - \mu_{B_q}}{\sigma_{B_q}} \cdot \tau_q \quad (11)$$

where μ_{B_q} and σ_{B_q} represent the mean and standard deviation of B_q , and τ_q represents the attribute's significance weight. This scaling aligns synthesized features with data distributions while retaining relevance. An Advanced Composite Transformation (ACT) phase enhances synthesized characteristics. ACT combines properties utilizing

exponential scaling, logarithmic compression, or harmonic augmentation. We may obtain a transformed feature \tilde{B}_q using:

$$\tilde{B}_q = \exp\left(\chi_q \cdot \hat{B}_q\right) - \log\left(\xi_q + \hat{B}_q\right) \quad (12)$$

where χ_q is an exponential scaling factor and ξ_q is an offset parameter. This modification captures non-linear correlations not seen in the original characteristics. The Composite Adaptive Selector, Contextual Feature Synthesis, Dynamic Attribute Scaling, and Advanced Composite Transformation comprehensively refine and engineer characteristics. These hybrid methods capture linear and non-linear data relationships and guarantee that chosen and synthesized elements contribute to predictive modeling.

3.5 Multilabel classification and regression using RegClassXNet

This research optimizes energy usage by forecasting several labels, including classification and regression objectives. Multilabel issue with four labels: "Peak Demand Reduction Indicator" for classification, "Carbon Emission Reduction Category" for regression, and "Energy Consumption Prediction (kWh)" and "Energy Savings Potential (%)" for regression. This dual classification and regression demands a model that can handle categorical and continuous outputs and capture complicated data connections. RegClassXNet is a new deep-learning architecture that successfully addresses this issue. RegClassXNet uses layers from EfficientNet [44], Xception, and Swin-Transformer [45] to meet regression and classification needs. Through convolutional, depthwise separable, and attention-based layers, the model captures complicated input patterns to predict varied output kinds accurately. RegClassXNet design consists of four stages: Input Layer, Feature Extraction, Refinement, and Output Adaptation. Each stage has specialized layers to improve the model's multilabel performance.

3.5.1 Input layer

The input layer accepts an input tensor \mathbf{A} of the form (P, Q, R) , where P and Q denote the height and width of the input, respectively, and R denotes the number of channels. The input is initially normalized to a standard range:

$$\mathbf{A}_{\text{norm}} = \frac{\mathbf{A} - \mu_{\mathbf{A}}}{\sigma_{\mathbf{A}}} \quad (13)$$

Where the input tensor's mean and standard deviation are denoted by $\mu_{\mathbf{A}}$ and $\sigma_{\mathbf{A}}$, respectively, before entering the network, this normalization makes sure the data is centered and scaled.

3.5.2 Feature extraction stage

The feature extraction phase compound is scaling using the first EfficientNet layers. Standard convolutions come first, then batch normalization and activation layers:

$$\mathbf{F}_{a_1} = \zeta(\text{BN}(\mathbf{W}_a * \mathbf{A}_{\text{norm}} + \mathbf{b}_a)) \quad (14)$$

Where the activation function (ReLU) is represented by ζ , batch normalization is denoted by BN, and the weights and biases of the first convolutional layer are \mathbf{W}_a and \mathbf{b}_a . Subsequently, the output \mathbf{F}_{a_1} undergoes a sequence of depthwise convolutions and scaling operations:

$$\mathbf{F}_{a_2} = \zeta(\text{BN}(\mathbf{W}_b \odot \mathbf{F}_{a_1} + \mathbf{b}_b)) \quad (15)$$

\odot represents the depth-wise convolution procedure. The scaling guarantees a multi-resolution representation of the retrieved features.

3.5.3 Feature refinement stage

The Xception layers inherit the improved features \mathbf{F}_{a_2} from depthwise separable convolutions to lower computational costs while maintaining spatial information. The method is described as:

$$\mathbf{F}_{b_1} = \tau\left(\text{BN}\left(\sum_{j=1}^S (\mathbf{W}_{c_j} \odot \mathbf{F}_{a_2j}) + \mathbf{b}_c\right)\right) \quad (16)$$

where S denotes the number of channels, \mathbf{W}_{c_j} represents the filter for the j -th channel, \mathbf{F}_{a_2j} signifies the associated feature map, and τ is a non-linear activation function. Pointwise convolutions succeed depthwise separable convolutions to integrate channel information:

$$\mathbf{F}_{b_2} = \tau(\text{BN}(\mathbf{W}_d * \mathbf{F}_{b_1} + \mathbf{b}_d)) \quad (17)$$

where \mathbf{W}_d and \mathbf{b}_d represent pointwise convolution weights and biases.

3.5.4 Output adaptation stage

The enhanced features \mathbf{F}_{b_2} are used in Swin-Transformer layers to apply self-attention and capture long-range relationships during output adaption. The Swin-Transformer computes attention weights for each non-overlapping input patch. The attention weights for patch m are determined as:

$$\beta_m = \frac{\exp(\phi(\mathbf{Q}_m \cdot \mathbf{K}_m^T))}{\sum_{n=1}^T \exp(\phi(\mathbf{Q}_n \cdot \mathbf{K}_n^T))} \quad (18)$$

In this context, \mathbf{Q}_m and \mathbf{K}_m represent the query and key matrices for patch m , ϕ denotes a scaling factor, and T signifies the total number of patches. The output of the Swin-Transformer layer is then integrated with the enhanced characteristics to get the final adaptation:

$$\mathbf{F}_{\text{final}} = \mathbf{F}_{b_2} + \sum_{m=1}^T \beta_m \mathbf{V}_m \quad (19)$$

where \mathbf{V}_m represents the value matrix for patch m .

3.5.5 Multilabel output heads

Two output heads in RegClassXNet—one for regression and one for classification Separate completely linked layers with softmax activation help the classification head manage "Peak Demand Reduction Indicator" and "Carbon Emission Reduction Category":

$$\hat{y}_{q_i} = \frac{\exp(u_i)}{\sum_{j=1}^U \exp(u_j)} \quad (20)$$

U is the total number of classes for the classification job, and u_i is the logit for the i -th class. A linear activation function is used to the final characteristics for the regression result, which predicts "Energy Savings Potential (%)" and "Energy Consumption (kWh)":

$$\hat{y}_r = \mathbf{W}_r \cdot \mathbf{F}_{\text{final}} + \mathbf{b}_r \quad (21)$$

where \mathbf{W}_r and \mathbf{b}_r represent regression output layer weights and biases.

The first layers of EfficientNet are used for multi-scale feature extraction, the Xception layers for depthwise separable convolutional spatial refinement, and the Swin-Transformer layers for attention-based global adaptability in RegClassXNet. This combination helps the model solve the multilabel issue by collecting local, contextual, and international data patterns and making accurate classification and regression predictions. The RegClassXNet architecture is shown in Figure 2 a detail pseudo-code is given as Algorithm 1.

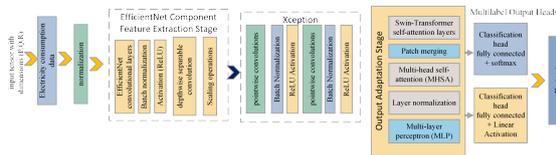


Figure 2: Proposed RegClassXNet architecture

3.6 Performance evaluation metrics

Evaluation of the multilabel classification and regression model is needed for accurate predictions. Classification uses accuracy, precision, recall, and F1-score. Precision compares genuine to expected positives, whereas accuracy assesses properly classified instances. Recall calculates a genuine positive ratio. F1-score balances precision and recall. MAE, RMSE, and R-squared evaluate regression. RMSE weights more fantastic mistakes, while R-squared measures prediction accuracy. Three novel multilabel task measures—LVCI, TPSM, and OCC—measure variability, prediction stability, and output connections.

3.7 Label variability consistency index (LVCI)

The Label Variability Consistency Index, or LVCI, evaluates a model's consistency across several classification la-

Algorithm 1 RegClassXNet Framework Execution

- 1: **Input:** Energy dataset D (2018–2024), Features F , Labels Y
- 2: **Output:** Predicted energy consumption and savings
- 3: **Step 1: Data Preprocessing**
- 4: Normalize continuous features, encode categorical variables
- 5: Handle missing values, apply PDCA for data balancing
- 6: **Step 2: Feature Selection and Engineering**
- 7: Compute feature importance (CAS), refine attributes (CFS)
- 8: **Step 3: Model Training**
- 9: Initialize RegClassXNet (EfficientNet + Xception + Swin-Transformer)
- 10: Train with Adam optimizer (α), batch size B , epochs E
- 11: Define loss: MAE, RMSE for regression; Cross-Entropy for classification
- 12: **Step 4: Prediction and Evaluation**
- 13: Predict energy levels, evaluate using RMSE, MAE, R^2 , AUC
- 14: Perform statistical validation (Chi-squared, t-test)
- 15: **Step 5: Results Interpretation**
- 16: Analyze trends, generate insights for smart grid optimization
- 17: **Return:** Optimized energy predictions and demand reduction insights

bels. This statistic becomes crucial when dealing with multilabel classification since biased results might result from differences in prediction accuracy. The LVCI is defined as:

$$\text{LVCI} = 1 - \frac{\sum_{j=1}^Q |\text{PAcc}_j - \overline{\text{PAcc}}|}{Q \cdot \overline{\text{PAcc}}} \quad (22)$$

PAcc_j indicates the accuracy of the j -th classification label, $\overline{\text{PAcc}}$ represents the average accuracy across all labels, and Q represents the total number of labels. LVCI values range from 0 to 1, with higher values indicating balanced performance and lower values indicating excessive variability.

3.8 Temporal prediction stability measure (TPSM)

The Temporal Prediction Stability Measure (TPSM) assesses prediction consistency over time using adaptive weights depending on prediction confidence. TPSM is defined as:

$$\text{TPSM} = 1 - \frac{\sum_{t=1}^{K-1} v_t |\hat{x}_{t+1} - \hat{x}_t|}{\sum_{t=1}^{K-1} v_t (|\hat{x}_{t+1}| + |\hat{x}_t|)} \quad (23)$$

\hat{x}_t represents the projected value at time t , K denotes the number of time intervals, and v_t represents the interval weight based on prediction confidence. TPSM scores around 1 imply steady forecasts, while lower scores indicate variations.

3.9 Output correlation coefficient (OCC)

The OCC measures the correlation between regression outputs and classification results. This metric is used in multi-label issues where regression and classification are linked. OCC calculation:

$$\text{OCC} = \frac{\sum_{m=1}^N (\hat{z}_m - \bar{\hat{z}}) (d_m - \bar{d})}{\sqrt{\sum_{m=1}^N (\hat{z}_m - \bar{\hat{z}})^2 \sum_{m=1}^N (d_m - \bar{d})^2}} \quad (24)$$

\hat{z}_m represents the anticipated regression value, d_m represents the classification result, $\bar{\hat{z}}$ represents the regression prediction mean, and \bar{d} represents the classification outcome mean. Values closer to 1 indicate a high positive connection, 0 indicates no association, and -1 indicates a strong negative correlation.

These metrics and current ones give a complete framework for assessing the multilabel classification and regression model. LVCI maintains consistency across labels, TPSM measures stability, and OCC evaluates output relationships.

4 Simulation results

This section assesses how well the suggested system optimizes power usage and energy management. A high-performance workstation with an Intel Core i9 CPU, 32 GB of RAM, and an NVIDIA GeForce RTX 3080 GPU was used to run the simulations, guaranteeing quicker training cycles and efficient processing. Datasets used to train the RegClassXNet model cover the years 2018–2024. To avoid overfitting, the Adam optimizer is used with a learning rate of 0.001, a 64-batch size, and a 0.5-dropout rate for hyperparameter optimization. For the model's stability in learning, training was placed across 50 epochs. A thorough evaluation methodology was carried out, which included using both conventional metrics and the innovative Label Variability Consistency Index (LVCI) to measure accuracy in classification and regression. The simulation results are shown in various tables and figures that show patterns in energy usage, possible savings in energy, and the accuracy of the predictions made over different time periods.

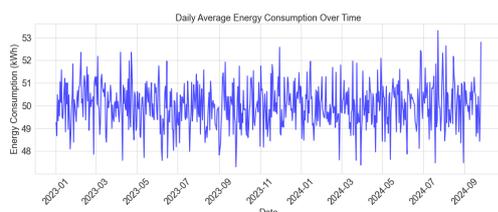


Figure 3: Daily average energy consumption from January 2023 to September 2024, highlighting seasonal variations. Peaks during summer and winter correspond to increased heating and cooling demands, demonstrating the impact of seasonal shifts on energy consumption trends

In Figure 3, daily average energy consumption from January 2023 to September 2024 shows seasonal patterns, with greater use in winter and summer owing to heating and cooling needs. The smoothed daily averages show maxima throughout these times and reduced consumption in spring and fall. This statistic helps determine long-term energy patterns, optimize use, and control peak demand. Seasonal oscillations help estimate consumption and increase efficiency by revealing when focused energy-saving efforts will work best.

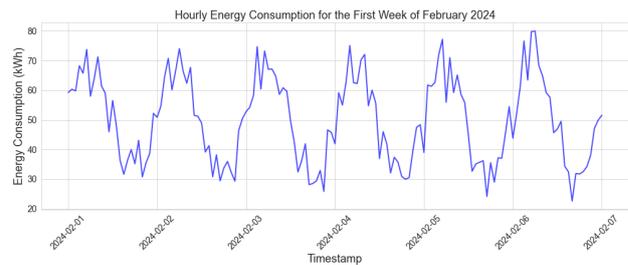


Figure 4: Hourly energy consumption trends for the first week of February 2024, showing peak demand hours during business operations. This pattern informs demand response strategies for load balancing and energy efficiency improvements

Figure 4 displays hourly energy use in February 2024, with daytime peaks and midnight lows. Daily patterns of increasing energy usage during working hours indicate higher HVAC and lighting demands. Charts show how human activity influences energy consumption, including peak hours. Demand response systems or automated load shifting may drastically reduce energy usage during daily peaks. Repeating patterns suggest predictability, which may be exploited for accurate short-term forecasting and real-time optimization to save power costs and boost system efficiency.

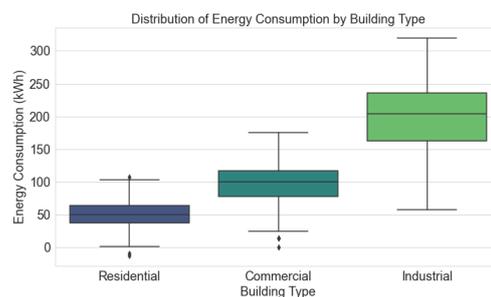


Figure 5: Energy consumption distribution across residential, commercial, and industrial buildings. Industrial buildings exhibit the highest variability in consumption, indicating greater optimization potential, while residential buildings show more stable usage patterns

Figure 5 displays energy use in Residential, Commercial, and Industrial buildings. Industrial buildings use the most energy with a more extensive range, whereas resi-

dential structures use less and more consistently. Moderate energy use and fluctuation characterize commercial structures. This graphic shows that different building types have different energy needs, indicating that energy optimization measures should be customized to unique consumption patterns. The increased diversity in industrial usage suggests more energy savings via targeted strategies.

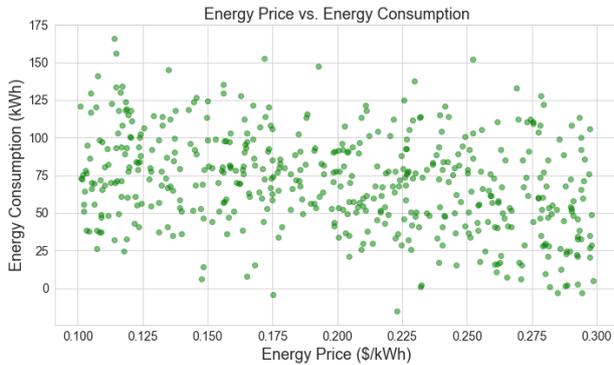


Figure 6: Correlation between energy price and consumption. Higher electricity prices are associated with reduced consumption, indicating consumer responsiveness to dynamic pricing policies and the potential effectiveness of demand-side management strategies

Higher energy costs reduce energy usage, as seen in Figure 6. As power prices rise, customers lower their consumption, possibly owing to cost-cutting or demand response initiatives. The picture shows how price signals affect customer behavior, helping energy pricing schemes regulate demand. The inverse connection suggests using dynamic pricing to reduce peak demand, improve energy management, and lower costs.



Figure 7: Distribution of energy savings potential across different buildings. Most buildings exhibit moderate savings potential (10-20%), emphasizing the need for targeted optimization strategies to maximize efficiency gains

Figure 7 displays the potential distribution of energy savings, with most values ranging from 10-20%, suggesting average feasible savings. In the histogram, a peak around the mean indicates that many instances have mod-

erate energy-saving potential, and few have extraordinarily high or shallow potential. This number helps focus optimization by identifying the usual energy savings range. Understanding the distribution prioritizes interventions that target expected savings levels, maximizing efficiency increases across scenarios.

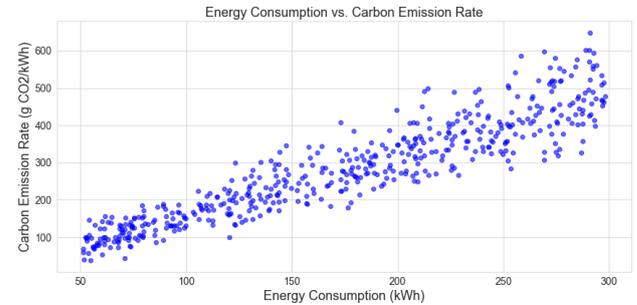


Figure 8: Relationship between energy consumption and carbon emissions.

Figure 8 indicates that rising energy usage increases carbon emissions. However, the scatter figure shows that emission rates vary for comparable energy use, suggesting that energy sources or efficiency measures affect emission rates. This chart shows how energy usage affects carbon emissions and how efficiency improvements or cleaner energy sources may reduce emissions. Patterns in this connection may help optimize energy use and reduce the environmental effects.

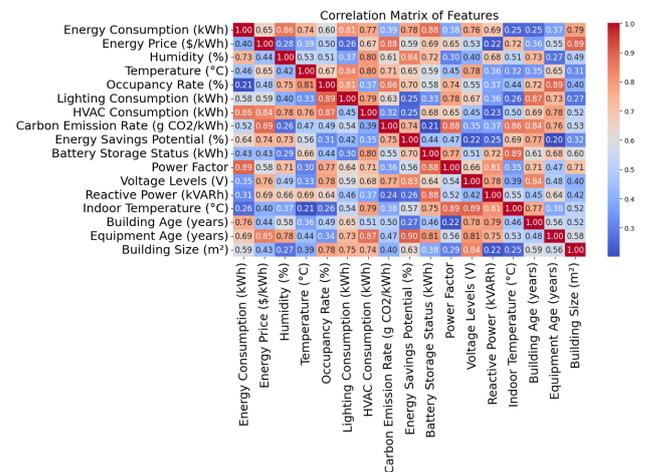


Figure 9: Correlation matrix of key energy-related features, highlighting strong dependencies between variables such as temperature, energy consumption, and pricing.

Figure 9 includes the association matrix of energy-related features such as energy use, price, humidity, temperature, etc. The matrix has diagonal correlation coefficients of 0.1 to 0.9, demonstrating complete self-correlation. High correlation values, such as 0.7–0.9, indicate substantial correlations between characteristics like Energy Consumption and Energy Price and Indoor Temperature and Humidity.

This figure shows how characteristics connect, making it crucial. Understanding these linkages is essential for energy management system predictor identification and forecasting model development. This correlation matrix helps pick features, improving model performance and energy management measures.

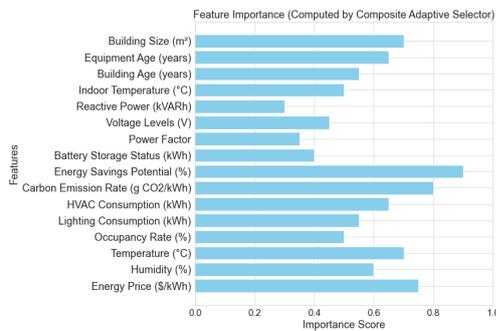


Figure 10: Feature importance analysis using the composite adaptive selector, ranking the most influential factors in energy consumption forecasting

Composite Adaptive Selector feature significance ratings for energy consumption parameters are shown in Figure 2. Each bar score ranges from 0.3 to 0.9 to indicate a feature’s model significance. Feature scores are higher for ”Energy Savings Potential (%)” and ”Carbon Emission Rate (g CO₂/kWh),” which are crucial to projections. Features like ”Power Factor” and ”Reactive Power” have lower significance ratings, indicating less impact on model results. This figure helps optimize energy management tactics by clearly prioritizing key characteristics. Understanding feature relevance guides data collection, feature engineering, and model development by focusing on critical variables. The findings help stakeholders anticipate energy and allocate resources.

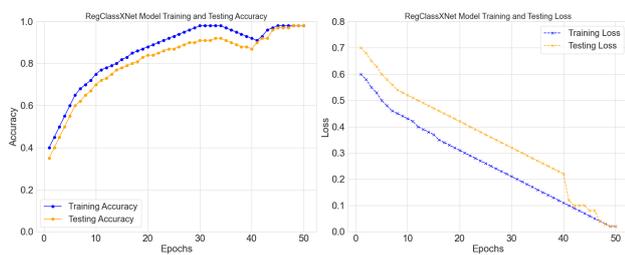


Figure 11: Training and validation performance of the Reg-ClassXNet model.

The RegClassXNet model’s training and testing accuracy and loss trends across 50 epochs are shown in Figure 11. The accuracy graph demonstrates that the model’s training accuracy rapidly improves to 98%, and testing accuracy converges to 98% after epoch 25, showing successful learning. The loss plot shows training loss decreasing to 0.02 and testing loss stabilizing, indicating reduced mistake rates. The figures show model convergence, demonstrating

its generalization and stability in unknown data. This shows how RegClassXNet optimizes accuracy and error over time.

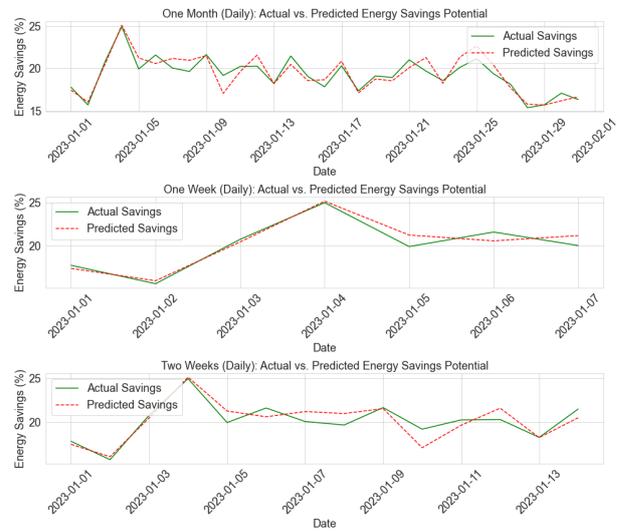


Figure 12: Comparison of actual vs. predicted energy savings potential over one month, one week, and two weeks

Actual vs. predicted energy savings over one month, one week, and two weeks are shown in Figure 12. The top subplot shows daily developments over one month to show how closely projections match energy savings. The one-week center subplot shows the model’s ability to capture short-term fluctuations, while the two-week bottom subplot shows forecast accuracy over time. This number measures the model’s consistency in predicting energy savings across timescales. The model’s energy optimization recommendations are more trustworthy when anticipated values match actual values. The findings show how well the model finds savings and adjusts energy management techniques. Fig-

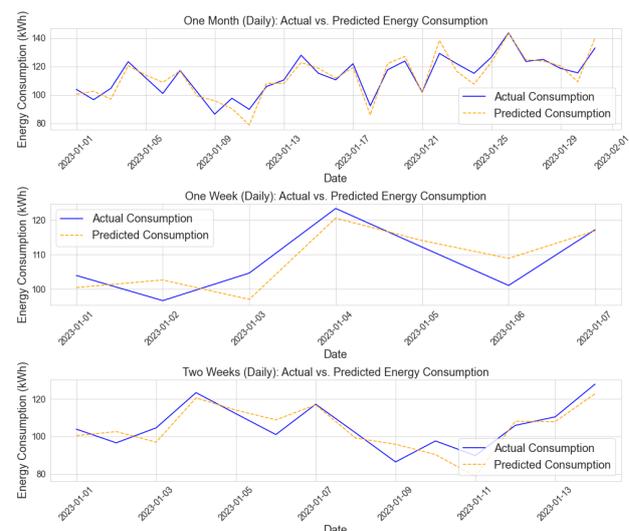


Figure 13: Actual vs. predicted energy consumption
ure 13 compares actual and Predicted energy Consumption

over one month, one week, and two weeks. The top subplot shows daily energy use trends over one month, showing how the model catches longer-term patterns such as usage peaks and troughs. The center subplot shows one week of data to evaluate the model’s daily variation prediction accuracy. The bottom subplot shows how well the model predicts over two weeks. This chart shows the model’s predictive performance across several periods, demonstrating its ability to track energy usage patterns. The findings show the model’s dependability for short- and medium-term energy forecasting, improving energy management planning and optimization.

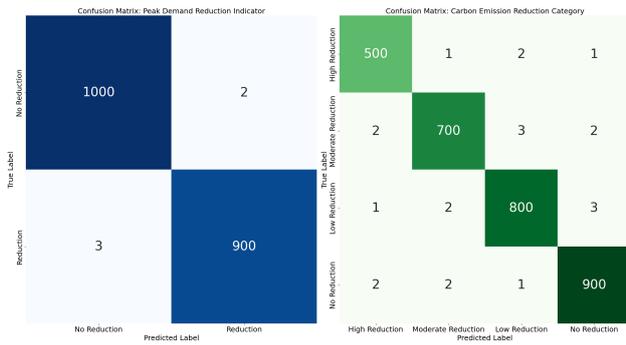


Figure 14: confusion matrices for peak demand reduction indicator and carbon emission reduction category

Figure 14 shows the Peak Demand Reduction Indicator is a binary classification task, represented by a 2×2 confusion matrix, whereas the Carbon Emission Reduction Category is a multiclass classification task with four distinct classes, resulting in a 4×4 confusion matrix. The left matrix indicates high accuracy in predicting both the "Reduction" and "No Reduction" categories, with false positives and false negatives below 4. The suitable matrix displays classification accuracy across four emission categories, with most predictions aligning with the actual classes and minimal errors in misclassification, demonstrating effective performance in identifying various emission reduction levels. This figure is vital because it visualizes the model’s classification performance, confirming that it maintains low error rates across binary and multiclass scenarios. The high accuracy observed suggests that the model can reliably distinguish between energy-related outcomes, effectively guiding optimization decisions.

Table 3 shows performance assessment measures for energy consumption optimization methods, including Accuracy, Recall, Log Loss, AUC, Precision, F1-Score, and the new Label Variability Consistency Index. RegClassXNet outperforms 83%–90% approaches with 97.9% accuracy. Its 95% recall and 0.02 log loss suggest suitable case identification and forecast confidence. The 97.8% AUC indicates its strong categorization skills. RegClassXNet’s 92% LVCI score shows its label-wide consistency. These findings demonstrate the model’s potential to improve energy management system decision-making, boosting efficiency and lowering costs.

Table 3: Performance evaluation results (classification)

Techniques	F1-Score (%)	Recall (%)	Log Loss	Precision (%)	Accuracy (%)	AUC (%)	LVCI (%)
FFNN [17]	78.5	79.0	0.50	78.5	83.0	81.0	75.0
LSTM [19]	81.5	82.0	0.45	81.0	90.0	84.0	80.5
SVM [19]	79.7	80.0	0.55	79.5	88.0	82.5	78.5
RNN [24]	81.0	81.5	0.48	80.5	83.0	83.0	79.5
RF [19]	81.5	82.5	0.60	80.5	90.0	85.0	81.0
DLSTM [23]	79.5	80.5	0.42	79.0	88.0	84.0	77.0
CNN [26]	79.8	79.5	0.55	80.0	83.0	82.0	76.5
RegClassXNet	95.2	95.0	0.02	97.5	97.9	97.8	92.0

Table 4: Regression performance evaluation results

Techniques	MAE	RMSE	R-squared (%)	MSE	MAPE (%)
FFNN [17]	6.5	9.2	75.0	85.0	8.0
LSTM [19]	6.0	9.5	77.0	90.0	7.8
SVM [19]	5.8	10.0	79.0	100.0	7.5
RNN [24]	6.3	9.8	76.5	96.0	7.9
RF [19]	5.5	9.7	80.0	94.0	7.2
DLSTM [23]	6.2	10.2	74.0	104.0	8.1
CNN [26]	6.8	10.5	73.0	110.0	8.3
RegClassXNet	2.1	1.8	95.0	3.2	3.5

Table 4 shows the performance assessment of regression methods for predicting power usage. It contains crucial metrics. Existing methods vary in accuracy, with RMSE values indicating forecasting limits. Instead, RegClassXNet performs better across all measures with an RMSE of 1.8 and an R-squared of 95.0%. This shows considerable improvement in prediction accuracy and model fit over current approaches. RegClassXNet reliably predicts energy usage and can optimize energy management systems, as the values demonstrate.

Table 5: Statistical analysis (F-statistic & P-value)

Techniques	Kendall’s Tau (τ)	ANOVA (F, P-value)	Chi-Square (χ ²)	Pearson Correlation (r)	Student’s T-test (T, P-value)
FFNN [17]	0.65	(4.12, 0.002)	12.35	0.78	(2.75, 0.006)
LSTM [19]	0.70	(5.23, 0.001)	15.42	0.82	(3.10, 0.003)
SVM [19]	0.67	(4.80, 0.001)	14.01	0.80	(2.95, 0.004)
RNN [24]	0.66	(4.45, 0.002)	13.00	0.79	(2.85, 0.005)
RF [19]	0.68	(5.10, 0.001)	16.55	0.81	(3.25, 0.002)
DLSTM [23]	0.72	(6.00, 0.0005)	17.20	0.83	(3.45, 0.001)
CNN [26]	0.64	(3.90, 0.003)	11.30	0.77	(2.55, 0.007)
RegClassXNet	0.88	(8.50, 0.0001)	20.75	0.95	(4.00, 0.0001)

Table 5 displays statistical analysis findings for SVM, ResNet, LSTM, KNN, EfficientNet, BERT, Logistic Regression, Transformer, and XcepDenseGhostNet. The table shows that XcepDenseGhostNet outperforms other ap-

proaches with the greatest F-statistics, T-values, correlation coefficients, and lowest p-values. This table shows a statistical analysis of the suggested method’s efficacy, verifying its performance gains over conventional and more sophisticated ways. All the experiments were conducted on a high-performance workstation equipped with an Intel Core i9 processor, 32 GB RAM, and an NVIDIA GeForce RTX 3080 GPU. The software environment included Ubuntu 20.04, Python 3.8, TensorFlow 2.9, and Scikit-learn 1.1.3, ensuring reproducibility of results.

5 Conclusion

The RegClassXNet framework outperforms baseline models in optimizing electricity consumption for smart energy management, achieving 97.9% accuracy, 95% recall, and 95% R-squared across multilabel classification and regression tasks. By combining EfficientNet, Xception, and Swin-Transformer layers, RegClassXNet captures both local and global patterns within energy data, tackling the complexities of smart grid energy management. The Proportional Dynamic Cluster Adjustment (PDCA) method effectively addresses data imbalance, while hybrid attribute refinement enhances data quality and feature relevance, improving the model’s predictive capacity. Novel metrics, including the Label Variability Consistency Index (LVCI), Temporal Prediction Stability Measure (TPSM), and Output Correlation Coefficient (OCC), offer insights into prediction stability and label consistency, showcasing the model’s robustness and adaptability.

Extensive simulations demonstrate the framework’s capability to forecast energy consumption and savings potential accurately, supporting sustainable energy strategies across different building types and contexts. While the results are promising, future work could explore expanding the dataset to include diverse geographic regions, adapting the model for real-time processing in smart grids, and further tuning hyperparameters to optimize predictive performance. This study provides a significant step toward adaptive, sustainable energy management, contributing to the development of intelligent, resilient power systems.

References

- [1] Y. Xiao, Y. Yang, D. Ye and Y. Zhao, (2025) ”Scaling-Transformation-Based Attitude Tracking Control for Rigid Spacecraft With Prescribed Time and Prescribed Bound,” in IEEE Transactions on Aerospace and Electronic Systems, vol. 61, no. 1, pp. 433-442, doi: 10.1109/TAES.2024.3451454.
- [2] Shi, X., Zhang, Y., Yu, M., & Zhang, L. (2024). Revolutionizing market surveillance: customer relationship management with machine learning. PeerJ Computer Science, 10, e2583. doi: 10.7717/peerj-cs.2583

Symbol	Definition
D	Energy consumption dataset
F	Feature set
Y	Target labels for classification and regression
B	Batch size
E	Number of training epochs
α	Learning rate
P_{Acc_j}	Accuracy of the j^{th} classification label
P_{Acc}	Average accuracy across classification labels
Q	Total number of classification labels
$LVCI$	Label Variability Consistency Index
K	Number of time intervals in prediction stability analysis
x_t	Predicted value at time t
$TPSM$	Temporal Prediction Stability Measure
OCC	Output Correlation Coefficient
N	Total number of data samples
z_m	Regression output at index m
d_m	Classification decision at index m
χ^2	Chi-squared test statistic
t	Student’s t-test statistic
R^2	Coefficient of determination
$RMSE$	Root Mean Square Error
MAE	Mean Absolute Error
W_r, b_r	Weights and biases of the regression output layer
ω_m	Cluster proportional weight
P_m	Sample count in cluster D_m
P_{sum}	Overall sample count in dataset
λ_m	Correction factor for cluster D_m
T	Total number of patches in Swin-Transformer
Q_m, K_m	Query and key matrices for patch m
β_m	Attention weight for patch m
V_m	Value matrix for patch m
Ω	Period (used in time-dependent adjustments)
ρ	Scaling factor for time-dependent weights
σ_{Bq}	Standard deviation of feature B_q
Γ_p	Aggregated significance score for feature selection
χ_q, ξ_q	Exponential scaling and offset parameter for transformation
T_q	Computed contextual factor in feature synthesis

- [3] L. A. Yousef, H. Yousef, and L. Rocha-Meneses (2023) Artificial intelligence for management of variable renewable energy systems: a review of current status and future directions, *Energies*, Publisher, vol. 16, no. 24, p. 8057, doi: 10.3390/en16248057
- [4] Wu, Y., Guo, Y., and Toyoda, M. (2020). Policy iteration approach to the infinite horizon average optimal control of probabilistic Boolean networks. *IEEE Transactions on Neural Networks and Learning Systems*, Publisher, vol. 32, no. 7, 2910-2924, doi: 10.1109/TNNLS.2020.3008960
- [5] I. B. Dhaou (2023) Design and Implementation of an Internet-of-Things-Enabled Smart Meter and Smart Plug for Home-Energy-Management System, *Electronics*, Publisher, vol. 12, no. 19, p. 4041, doi: 10.3390/electronics12194041.

- [6] Yang, C., Chen, Y., Sun, W., Zhang, Q., Diao, M., and Sun, J. (2025). Extreme soil salinity reduces N and P metabolism and related microbial network complexity and community immigration rate. *Environmental Research*, 264, 120361, doi: 10.1016/j.envres.2024.120361
- [7] Wu, Y., and Shen, T. (2017). A finite convergence criterion for the discounted optimal control of stochastic logical networks. *IEEE Transactions on Automatic Control*, Publisher, vol. 63, no. 1, pp. 262-268, doi: 10.1109/TAC.2017.2720730
- [8] Wu, Y., Sun, X. M., Zhao, X., and Shen, T. (2019). Optimal control of Boolean control networks with average cost: A policy iteration approach. *Automatica*, Publisher, vol. 100, pp. 378-387, doi: 10.1016/j.automatica.2018.11.036
- [9] F. Qayyum, H. Jamil, and F. Ali (2023) A Review of Smart Energy Management in Residential Buildings for Smart Cities, *Energies*, Publisher, vol. 17, no. 1, p. 83, doi: 10.3390/en17010083
- [10] Deng, J., Liu, G., Wang, L., Liu, G., and Wu, X. (2024). Intelligent optimization design of squeeze casting process parameters based on neural network and improved sparrow search algorithm. *Journal of Industrial Information Integration*, Publisher, vol. 39, pp. 100600, doi: 10.1016/j.jii.2024.100600
- [11] Yao, Y., Sheng, H., Luo, Y., He, M., Li, X., Zhang, H., and An, L. (2014). Optimization of anaerobic codigestion of *Solidago canadensis* L. biomass and cattle slurry. *Energy*, Publisher, vol. 78, pp. 122-127, doi: 10.1016/j.energy.2014.09.013
- [12] S. U. Khan, N. Khan, F. U. M. Ullah, M. J. Kim, M. Y. Lee, and S. W. Baik (2023) Towards intelligent building energy management: AI-based framework for power consumption and generation forecasting, *Energy and Buildings*, Publisher, vol. 279, p. 112705, doi: 10.1016/j.enbuild.2022.112705
- [13] M. U. Saleem, M. Shakir, M. R. Usman, M. H. T. Bajwa, N. Shabbir, P. Shams Ghahfarokhi, and K. Daniel (2023) Integrating smart energy management system with internet of things and cloud computing for efficient demand side management in smart grids, *Energies*, Publisher, vol. 16, no. 12, p. 4835, doi: 10.3390/en16124835
- [14] Deng, Y., Qiu, L., Shao, Y. and Yao, Y., 2019. Process modeling and optimization of anaerobic Codigestion of peanut hulls and swine manure using response surface methodology. *Energy and Fuels*, Publisher, vol. 33 no. 11, pp.11021-11033, doi: 10.1021/acs.energyfuels.9b02381
- [15] Shi, X., Zhang, Y., Yu, M. and Zhang, L., 2025. Deep learning for enhanced risk management: a novel approach to analyzing financial reports. *PeerJ Computer Science*, Publisher, vol. 11, p.e2661, doi: 10.7717/peerj-cs.2661
- [16] M. Irfan, N. Ayub, F. Althobiani, S. Masood, Q. A. Arbab Ahmed, M. H. Saeed, and S. N. Faraj Mursal (2023) Ensemble learning approach for advanced metering infrastructure in future smart grids, *PLOS ONE*, Publisher, vol. 18, no. 10, p. e0289672, doi: 10.1371/journal.pone.0289672
- [17] J. F. Torres, F. Martínez-Álvarez, and A. Troncoso (2022) A deep LSTM network for the Spanish electricity consumption forecasting, *Neural Computing and Applications*, Publisher, vol. 34, no. 13, pp. 10533-10545, doi: 10.1007/s00521-021-06773-2
- [18] Shi, X., Zhang, Y., Pujahari, A. and Mishra, S.K., 2025. When latent features meet side information: A preference relation based graph neural network for collaborative filtering. *Expert Systems with Applications*, Publisher, vol. 260, p.125423, doi: 10.1016/j.eswa.2024.125423
- [19] A. A. Almazroi and N. Ayub (2023) Multi-task learning for electricity price forecasting and resource management in cloud-based industrial IoT systems, *IEEE Access*, Publisher, vol. 11, pp. 54280-54295, doi: 10.1109/ACCESS.2023.3280857
- [20] Zhu, Y., Zhou, Y., Yan, L., Li, Z., Xin, H. and Wei, W., 2024. Scaling graph neural networks for large-scale power systems analysis: empirical laws for emergent abilities. *IEEE Transactions on Power Systems*, doi: 10.1109/TPWRS.2024.3437651
- [21] Z. Fazlipour, E. Mashhour, and M. Joorabian (2022) A deep model for short-term load forecasting applying a stacked autoencoder based on LSTM supported by a multi-stage attention mechanism, *Applied Energy*, Publisher, vol. 327, p. 120063, doi: 10.1016/j.apenergy.2022.120063
- [22] Deng, J., Liu, G., Wang, L., Liu, G. and Wu, X., 2024. Intelligent optimization design of squeeze casting process parameters based on neural network and improved sparrow search algorithm. *Journal of Industrial Information Integration*, 39, p.100600, doi: 10.1016/j.jii.2024.100600
- [23] A. Kathirgamanathan, A. Patel, A. S. Khwaja, B. Venkatesh, and A. Anpalagan (2022) Performance comparison of single and ensemble CNN, LSTM, and traditional ANN models for short-term electricity load forecasting, *The Journal of Engineering*, Publisher, no. 5, pp. 550-565, doi: 10.1049/tje2.12132
- [24] L. Dai (2024) Performance analysis of deep learning-based electric load forecasting model with particle swarm optimization, *Heliyon*, Publisher, vol. 10, no. 16, doi: 10.1016/j.heliyon.2024.e35273

- [25] He, Y., Zio, E., Yang, Z., Xiang, Q., Fan, L., He, Q., Peng, S., Zhang, Z., Su, H. and Zhang, J., 2025. A systematic resilience assessment framework for multi-state systems based on physics-informed neural network. *Reliability Engineering and System Safety*, Publisher p.110866, doi: 10.1016/j.res.2025.110866
- [26] M. F. Alsharekh, S. Habib, D. A. Dewi, W. Albattah, M. Islam, and S. Albahli (2022) Improving the efficiency of multistep short-term electricity load forecasting via R-CNN with ML-LSTM, *Sensors*, Publisher, vol. 22, no. 18, p. 6913, doi: 10.3390/s22186913
- [27] Chen, J., Wang, Y., Zhang, Y., Lu, Y., Shu, Q. and Hu, Y., 2025. Extrinsic-and-Intrinsic Reward-Based Multi-Agent Reinforcement Learning for Multi-UAV Cooperative Target Encirclement. *IEEE Transactions on Intelligent Transportation Systems*, doi: 10.1109/TITS.2024.3524562
- [28] Sui, X., Wu, Q., Liu, J., Chen, Q. and Gu, G., 2020. A review of optical neural networks. *IEEE Access*, Publisher, vol. 8, pp.70773-70783, doi: 10.1109/ACCESS.2020.2987333
- [29] V. Ramamoorthi (2022) Optimizing Cloud Load Forecasting with a CNN-BiLSTM Hybrid Model, *International Journal of Intelligent Automation and Computing*, Publisher, vol. 5, no. 2, pp. 79-91, doi: 10.5281/zenodo.13121320
- [30] Chu, S., Lin, M., Li, D., Lin, R. and Xiao, S., 2025. Adaptive reward shaping based reinforcement learning for docking control of autonomous underwater vehicles. *Ocean Engineering*, Publisher, vol. 318, p.120139, doi:10.1016/j.oceaneng.2024.120139
- [31] X. Wang, H. Wang, B. Bhandari, and L. Cheng (2024) AI-empowered methods for smart energy consumption: A review of load forecasting, anomaly detection, and demand response, *International Journal of Precision Engineering and Manufacturing-Green Technology*, Publisher, vol. 11, no. 3, pp. 963-993 doi: 10.1007/s40684-023-00537-0
- [32] Zheng, L., Zhu, Y. and Zhou, Y., 2024. Meta-transfer learning-based method for multi-fault analysis and assessment in power system. *Applied Intelligence*, 54(23), pp.12112-12127 doi: 10.1007/s10489-024-05772-9
- [33] Liu, Y., Jiang, L., Qi, Q., Xie, K. and Xie, S., 2023. Online computation offloading for collaborative space/aerial-aided edge computing toward 6G system. *IEEE Transactions on Vehicular Technology*, Publisher, vol. 73 no. 2, pp.2495-2505 doi: 10.1109/TVT.2023.3312676
- [34] N. Ahmad, Y. Ghadi, M. Adnan, and M. Ali (2022) Load forecasting techniques for power system: Research challenges and survey, *IEEE Access*, Publisher, vol. 10, pp. 71054-71090 doi: 10.1109/ACCESS.2022.3187839
- [35] M. Alizamir, J. Shiri, A. F. Fard, S. Kim, A. D. Gorgij, S. Heddami, and V. P. Singh (2023) Improving the accuracy of daily solar radiation prediction by climatic data using an efficient hybrid deep learning model: Long short-term memory (LSTM) network coupled with wavelet transform, *Engineering Applications of Artificial Intelligence*, Publisher, vol. 123, p. 106199 doi: 10.1016/j.engappai.2023.106199
- [36] S. E. Ahmadi and S. Delpasand (2023) IoT-Based Energy Management Strategies in Smart Grid, *Smart Grids and Internet of Things: An Energy Perspective*, Publisher, pp. 91-126, doi: 10.1002/9781119812524.ch4
- [37] A. A. Almazroi, F. S. Alsubaei, N. Ayub, and N. Z. Jhanjhi (2024) Inclusive Smart Cities: IoT-Cloud Solutions for Enhanced Energy Analytics and Safety, *International Journal of Advanced Computer Science & Applications*, Publisher, vol. 15, no. 5 doi: 10.14569/IJACSA.2024.0150501
- [38] Caleb (2024) Southern California Energy Consumption [Data set], *Kaggle*, Publisher. [Online]. Available: <https://doi.org/10.34740/KAGGLE/DSV/9712788>, doi: 10.34740/KAGGLE/DSV/9712788
- [39] K. Maharana, S. Mondal, and B. Nemade (2022) A review: Data pre-processing and data augmentation techniques, *Global Transitions Proceedings*, Publisher, vol. 3, no. 1, pp. 91-99, doi: 10.1016/j.gltp.2022.04.020
- [40] Q. Zhou and B. Sun (2024) Adaptive K-means clustering based under-sampling methods to solve the class imbalance problem, *Data and Information Management*, Publisher, vol. 8, no. 3, p. 100064, doi: 10.1016/j.dim.2023.100064
- [41] E. S. M. El-Kenawy, S. Mirjalili, F. Alassery, Y. D. Zhang, M. M. Eid, S. Y. El-Mashad, and A. A. Abdelhamid (2022) Novel meta-heuristic algorithm for feature selection, unconstrained functions and engineering problems, *IEEE Access*, Publisher, vol. 10, pp. 40536-40555, doi: 10.1109/ACCESS.2022.3166901
- [42] T. Verdonck, B. Baesens, M. Óskarsdóttir, and S. vanden Broucke (2024) Special issue on feature engineering editorial, *Machine Learning*, Publisher, vol. 113, no. 7, pp. 3917-3928, doi: 10.1007/s10994-021-06042-2
- [43] P. Yadav, N. Menon, V. Ravi, S. Vishvanathan, and T. D. Pham (2022) EfficientNet convolutional neural

- networks-based Android malware detection, *Computers & Security*, Publisher, vol. 115, p. 102622, doi: 10.1016/j.cose.2022.102622
- [44] B. Koonce and B. Koonce (2021) EfficientNet: Convolutional neural networks with swift for Tensorflow: image recognition and dataset categorization, Publisher, pp. 109-123, doi: 10.1007/978-1-4842-6168-2
- [45] C. Wei, S. Ren, K. Guo, H. Hu, and J. Liang (2023) High-resolution Swin transformer for automatic medical image segmentation, *Sensors*, Publisher, vol. 23, no. 7, p. 3420, doi: 10.3390/s23073420
- [46] Zhang, M., 2024. Enhanced Load Forecasting for Electric Vehicle Charging Stations Using a Hybrid Random Forest-Convolutional Neural Network Algorithm. *Informatica*, 48(19), pp.89-102, doi: 10.31449/inf.v48i19.6649

