Enhanced IoT-Driven Load Forecasting with Metaheuristic-Optimized Deep Learning for Logistics Planning

Ramadan Babers[1], Walid Atwa[2], Mohamed Meselhy Eltoukhy[3], A. M. M. Madbouly[4]
Faculty of Computer Studies, Arab Open University, Cairo, Egypt[1]
Computer Science Department-Faculty of Computers and Information, Menoufia University, Menoufia, Egypt[2]
Department of Information Technology-College of Computing and Information Technology at Khulais,
University of Jeddah, Jeddah, Saudi Arabia[3]
Mathematics Department-Faculty of Science, Capital (Helwan) University, Cairo, Egypt[4]

Abstract—**The integration of IoT technologies with smart logistics operations has opened unprecedented avenues for optimizing energy consumption in warehouse facilities. Accurate forecasting of electricity load is a key factor in cost reduction, operational efficiency, and sustainable energy management. This study presents an Enhanced Integrated Load Forecasting System (E-ILFS) that synergizes metaheuristic optimization with deep learning architectures of higher order for superior electricity load forecasting in dynamic logistics environments. Building on the foundational ILFS framework, our enhanced approach integrates Harris Hawks Optimization (HHO) for robust feature selection and an improved Residual Network (ResNet) enhanced with self-supervised learning (SSL) to more effectively capture complex, non-linear temporal patterns. Finally, comprehensive experimental evaluation on a real-world IoT-driven logistics dataset demonstrates that E-ILFS achieves state-of-the-art performance with an R² score of 0.8745, MAE of 23.59, and MAPE of 3.22%, representing a significant 12.51% improvement in R² over baseline models. In fact, the proposed system provides a practical and scalable solution for real-world logistics operations.**

Keywords*—*Electricity load forecasting, IoT; logistics planning; metaheuristic optimization; deep learning; Harris Hawks optimization; ResNet; self-supervised learning

I. INTRODUCTION

The rapid expansion of global supply chains and growing urbanization have dramatically intensified energy demands in logistics operations. Thus, electricity consumption behaviors show complex patterns driven by shipping schedules, fluctuating occupancy, seasonal demand, and operational metrics in warehouse facilities. Moreover, it is challenging for traditional forecasting models to capture the intricate nonlinear relationships around this data-rich environment. IoT technologies enable the collection of real-time data, opening opportunities for sophisticated, data-driven forecasting approaches.

The Integrated Load Forecasting System (ILFS) introduced a hybrid framework combining Boruta-XGBoost for feature selection with Hybrid Huber Regression and ResNet (HRRN) [1]. While effective, its performance leaves room for improvement in feature-selection robustness and deep-learning architecture refinement.

To address these challenges, prior work introduced the Integrated Load Forecasting System (ILFS), which represented an important step toward a unified, data-driven framework for electricity load forecasting in IoT-equipped logistics settings [1]. The ILFS framework adopted a two-stage architecture. In the first stage, a Boruta-XGBoost feature-selection pipeline was applied to identify the most predictive variables from the high-dimensional IoT sensor data, leveraging the complementary strengths of Boruta's shadow-feature-based variable importance assessment and XGBoost's gradient-boosted ensemble learning for robust feature ranking [2]. In the second stage, the HRRN model was trained on the selected feature subset to produce electricity load forecasts, combining the robustness of Huber loss minimization against outlier contamination with the representational power of residual network architectures for non-linear temporal pattern learning [1]. The ILFS demonstrated meaningful improvements over purely statistical baselines, validating the core hypothesis that the integration of intelligent feature selection with deep learning-based forecasting can yield superior predictive accuracy in logistics energy management. Nevertheless, a careful examination of the ILFS architecture and its empirical performance reveals several important limitations that constrain its effectiveness and generalizability [3].

First, although efficient for moderately sized datasets, the Boruta-XGBoost feature selection mechanism exhibits limitations in the extremely high-dimensional and non-stationary feature spaces characteristic of large-scale IoT setups [4]. Boruta is an inherently wrapper feature selection method that depends on an iterative permutation test involving shadow features. Consequently, this process becomes computationally prohibitive as the number of sensor channels scales into the hundreds or thousands; it also exhibits unstable performance when the data is characterized by high levels of temporal autocorrelation [4]. The feature importance scores computed using XGBoost are found to be cardinality dependent and tend to have biased results, where the importance of continuous features is underestimated compared to categorical ones, leading to suboptimal feature sets in the case of IoT data, which is heterogeneous in nature [5].

Second, while the ResNet portion of the HRRN framework itself is a legitimate approach to extracting hierarchical features from time series data, its implementation in the ILFS framework does not take advantage of the large quantities of unlabeled or weakly labeled data, such as those commonly encountered in IoT-based logistics settings. Indeed, collecting and preparing precisely labeled data related to electricity consumption during a variety of operational activities can be labor-intensive and, in many cases, result in a dataset with a relatively small size, leading to overfitting and a lack of ability to generalize to new operational conditions [6] The inability of the original ILFS framework to implement self-supervised pre-training and contrastive learning means that the model does not take advantage of large quantities of unlabeled time series data, such as those provided by IoT sensors during periods when ground truth energy allocation labels are not provided.

Third, the ILFS does not incorporate any explicit mechanism for hyperparameter optimization of the deep learning model, relying instead on manually tuned configurations. The performance of deep ResNet architectures is known to be highly sensitive to hyperparameter choices including learning rate schedules, depth, filter sizes, dropout rates, and regularization coefficients. Without systematic, principled optimization of these parameters, the ILFS ResNet may fail to realize its full representational potential, leaving significant predictive accuracy on the table [7].

Fourth, empirical evaluations of the ILFS have revealed that its performance, while competitive, leaves room for meaningful improvement across standard regression metrics. Specifically, the baseline R^2 scores, mean absolute errors (MAE), and mean absolute percentage errors (MAPE) reported for the ILFS indicate that the system occasionally struggles to capture sharp load fluctuations associated with irregular logistics events and to generalize across facilities with different operational profiles, suggesting that both the feature-selection and learning stages require enhancement to achieve the reliability required for real-world deployment at scale [1].

To address the limitations, we employ Harris Hawks Optimization (HHO) [8] for robust feature selection in high-dimensional IoT data and an improved ResNet [9] with self-supervised learning capabilities to enhance pattern recognition. Comprehensive empirical evaluations demonstrate performance gains of up to 12.51% in R^2 and reductions of up to 24.35% in MAE compared with baseline models.

Motivated by the limitations, this study presents the Enhanced Integrated Load Forecasting System (E-ILFS), a comprehensive redesign of the ILFS that addresses each identified deficiency through principled methodological advances. E-ILFS introduces two principal innovations relative to its predecessor.

The first innovation concerns feature selection. In place of the Boruta-XGBoost pipeline, E-ILFS employs HHO, a nature-inspired metaheuristic algorithm that emulates the cooperative predatory behavior of Harris's hawk populations [10]. HHO has demonstrated superior performance relative to established optimization methods across a wide range of feature-selection benchmarks, particularly in high-dimensional, noisy search spaces that mirror the characteristics of IoT sensor data. By formulating feature subset selection as a combinatorial optimization problem and applying HHO to search the exponentially large space of possible feature subsets guided by a prediction-accuracy fitness function, E-ILFS can identify more compact and informative feature sets compared to the wrapper-based Boruta approach, reducing the risk of selecting redundant or noise-amplifying features while lowering the computational overhead of the selection process for large-scale IoT deployments [11].

The second innovation concerns the deep learning forecasting architecture. E-ILFS replaces the standard ResNet component of the HRRN with an improved ResNet augmented by a self-supervised learning (SSL) pre-training stage. Self-supervised learning enables the model to learn rich, generalizable representations of temporal patterns from large quantities of unlabeled IoT data that is routinely available in logistics environments but would be discarded by supervised-only training pipelines. Specifically, E-ILFS adopts a masked temporal modeling pre-training objective, analogous to masked language modeling in the natural language processing domain, in which the model learns to reconstruct deliberately obscured segments of the input time series [12]. This pre-training instills the model with a strong inductive bias toward the underlying temporal dynamics of energy consumption before fine-tuning on the labeled forecasting task, substantially improving generalization and robustness to distributional shifts caused by changes in logistics operations.

Together, these innovations produce a system that is more accurate, more computationally efficient in its feature-selection stage, and better able to exploit the full informational content of IoT-generated data streams.

The remainder of this study is structured as follows. Section II provides a comprehensive review of related work, covering IoT-based energy management in logistics, traditional and deep learning approaches to electricity load forecasting, feature-selection methods for high-dimensional sensor data, and metaheuristic optimization algorithms with particular focus on HHO. Section III describes the problem formulation and the architecture of the proposed E-ILFS framework, detailing the HHO-based feature-selection pipeline, the SSL-augmented ResNet design, and the integration mechanism that connects these components into a coherent end-to-end forecasting system. Section IV presents the experimental setup, including a description of the real-world IoT logistics dataset, the evaluation protocol, the competing baseline models, and the performance metrics adopted. Section V concludes the study and suggests directions for future work.

II. RELATED WORK

Accurate electricity load forecasting is fundamental to intelligent energy management, particularly in logistics, where operational variability and energy costs intersect. Existing approaches, spanning statistical, machine learning, and deep learning paradigms, offer complementary strengths but vary in their ability to effectively handle the high-dimensional data generated by IoT-enabled logistics environments.

Classical time-series models, such as Autoregressive Integrated Moving Average (ARIMA) and its seasonal variant SARIMA [13], have been widely used for load forecasting due to their mathematical tractability and strong performance on stationary data with linear trends and seasonal patterns. However, their inherent linearity limits their ability to model the complex, nonlinear dynamics of electricity consumption logistics, which is influenced by volatile factors such as shipping activity, weather conditions, and real-time warehouse operations. Moreover, their limited capacity to natively incorporate exogenous variables, including temperature and occupancy, restricts their effectiveness in IoT-driven, multivariate forecasting environments.

The emergence of machine learning (ML) techniques introduced greater flexibility in overcoming the limitations of statistical models, with methods such as Support Vector Regression (SVR) [14] and ensemble approaches like Random Forests (RF) [15]. [1] demonstrated improved capability in modeling nonlinear relationships. These models can incorporate diverse feature sets, leading to enhanced performance in complex environments. However, their effectiveness depends heavily on manually engineered features, domain expertise, and careful hyperparameter tuning. Moreover, they struggle to capture long-range temporal dependencies and hierarchical patterns in sequential data, which are essential for modern forecasting tasks involving multi-scale periodicity (e.g., hourly, daily, weekly, and seasonal cycles).

The rise of deep learning has radically changed the landscape for sequence forecasting. Architectures such as Long Short-Term Memory (LSTM) networks do exceptionally well in learning long-term temporal dependencies. CNNs, generally applied to spatial data, have been modified to capture local temporal patterns through 1D convolutions.

Hybrid models, such as CNN-LSTM frameworks [16], combine the strengths of both, using CNNs for local feature extraction and LSTMs for temporal sequence modeling. Very recently, Residual Networks (ResNet) [9], [17], while conceived for computer vision, have been recycled for time-series forecasting. Their skip connection architecture avoids the vanishing gradient problem, thus allowing the training of much deeper networks, capable of learning highly complex, non-linear mappings. Despite these advances, pure supervised deep learning models often require large quantities of labeled data and may not learn robust representations when faced with noise and incompleteness in real-world IoT data streams.

Zhang et al. [18] designed a real-time load forecasting model using a CNN-BiLSTM framework where Bayesian optimization is used to optimize learning parameters such as the learning rate and batch size. In their methodology, spatial-temporal features are extracted using a 1D CNN followed by capturing the bidirectional time dependency using a BiLSTM model. Bayesian optimization is employed for parameter tuning in this framework. The performance of their model has been evaluated on various publicly available datasets. Their proposed model gives an accuracy of 0.963 and an area under the curve (AUC) score of 0.951 and hence outperforms the traditional approaches like ARIMA, support vector machine (SVM), and LSTM. However, the proposed framework does not perform feature selection, nor does it consider self-supervised learning, thus highlighting the need for our proposed E-ILFS technique for logistics forecasting applications.

Feature selection is a critical preprocessing step in high-dimensional forecasting tasks, and metaheuristic algorithms offer an effective, gradient-free means of exploring complex search spaces. Among them, HHO has gained attention for its efficiency, inspired by the cooperative hunting strategy of Harris hawks and its balanced exploration–exploitation mechanism. This makes HHO well suited for identifying optimal feature subsets from large IoT datasets. However, its use in optimizing the input space of deep learning models remains underexplored, presenting substantial potential for improving model efficiency and predictive accuracy.

Despite steady progress in the literature, a clear gap remains in the joint integration of robust metaheuristic feature selection and advanced self-supervised deep learning architectures for logistics load forecasting, as most studies address these components in isolation. Moreover, the application of self-supervised learning (SSL) to enrich temporal representations from unlabeled data is still in its early stages within this domain. To bridge these gaps, this study proposes the Enhanced Integrated Load Forecasting System (E-ILFS).

The contributions are threefold: 1) the first integration of HHO for intelligent feature selection in high-dimensional logistics IoT data, 2) an enhanced ResNet forecasting core augmented with SSL for improved temporal representation learning, and 3) extensive empirical validation demonstrating substantial performance gains over established baselines and recent state-of-the-art methods.

III. PROPOSED ENHANCED SYSTEM (E-ILFS)

The Enhanced Integrated Load Forecasting System (E-ILFS) is an end-to-end system that converts IoT raw, multivariate data into an accurate and reliable load forecasting output. The system architecture, as presented in Fig. 1, comprises four major components that are integrated to deliver high-quality output. The components are: 1) Data Preprocessing & Feature Engineering, which preprocesses the IoT data to make it suitable for forecasting purposes, 2) Metaheuristic Feature Selection, which utilizes the HHO approach to determine the best feature subset for forecasting purposes, 3) Deep Learning Forecasting, which utilizes an enhanced ResNet model with self-supervised learning (SSL) and LSTM to deliver an accurate output, and 4) Optimization Module, which optimizes the performance of the system and ensures the forecasting output is delivered efficiently.

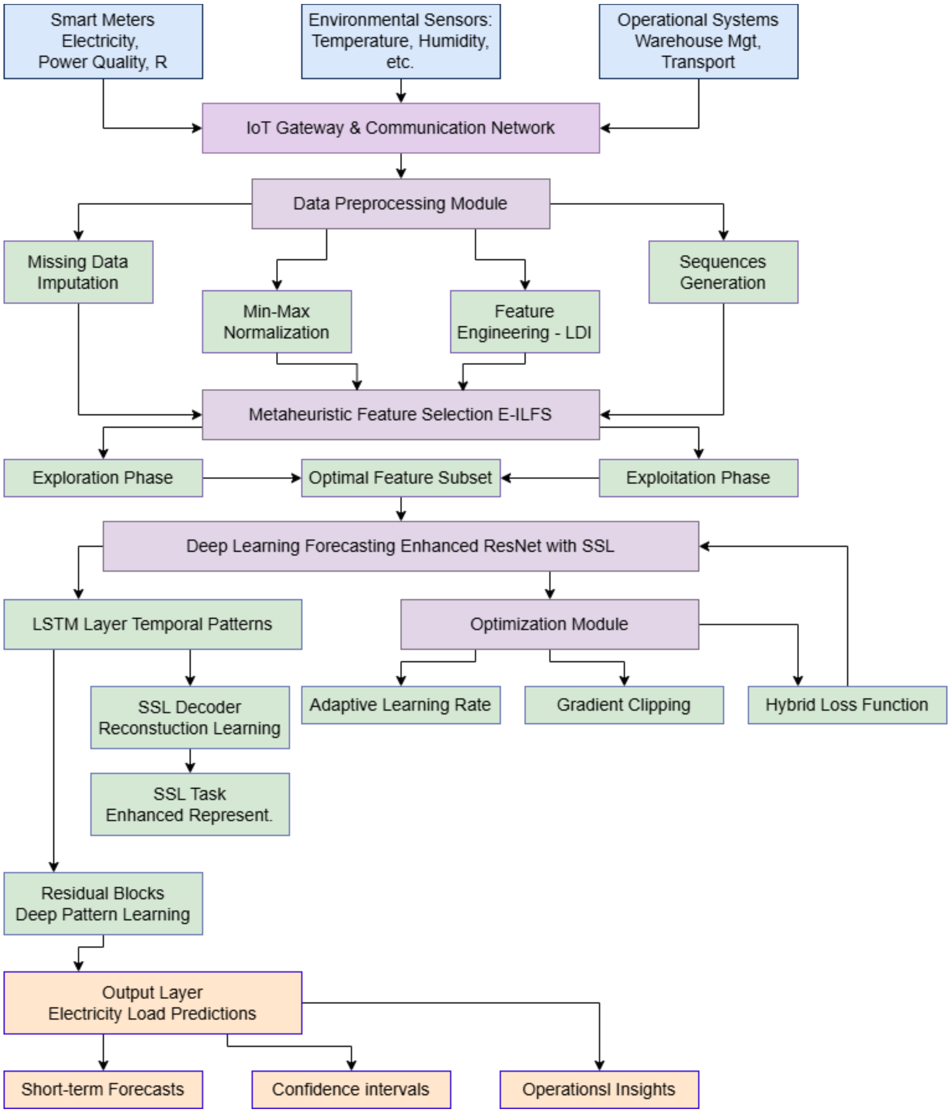

Fig. 1. Architecture of the proposed enhanced integrated load forecasting system (E-ILFS).

A. Data Collection and Preprocessing

The research uses the publicly available Electricity Load Logistics IoT Dataset, available online and spanning the years 2015-2024, containing 3317 instances, each recorded hourly [19]. The dataset includes various characteristics of logistics-related energy consumption, which are multidimensional, including: Temporal Features: Hour of the day, day of the week, holiday flags. Environmental Factors: Ambient temperature, humidity. Operational Metrics: Occupancy rate, equipment utilization, transportation levels. Energy Market Data: Day-ahead and real-time energy prices (Commodity Charge – CC, Marginal Loss Charge – MLC), system load, regulation capacity. A robust preprocessing process guarantees the quality of the data and the model's readiness. Missing values are handled by the k-NN algorithm with k = 5. All numerical values are normalized to the range [0,1] using the Min-Max normalization technique to improve numerical stability and speed up the convergence of the neural network model. Moreover, a new composite feature called the Load Density Index is created by combining the occupancy levels with the scheduled shipment information. This new feature directly corresponds to the energy demand and increases the predictive value of the dataset.

B. Improved Feature Selection Using Harris Hawks Optimization (HHO)

The preprocessed data is initially provided with 23 features, and this may cause overfitting and may also require a high computational cost. To mitigate this, an intelligent feature selection approach, known as Harris Hawks Optimization (HHO), is used. The HHO is a mathematical imitation of the cooperative behavior of Harris hawks while hunting, including surprise pounce and chase, and is divided into three phases: exploration, transition, and exploitation. The exploration phase is a global search, while the transition and exploitation phases are local searches. The global search is highly efficient compared to other methods. The feature-selection problem is formulated as a binary optimization task. Each candidate solution represents a subset of features, and the fitness function is defined as:

$$\text{Fitness} = \alpha \cdot \text{MAE}_{\text{model}} + \beta \cdot \frac{|S|}{|F|}$$

where, $\text{MAE}_{\text{model}}$ is the mean absolute error of a simple predictor (e.g., a shallow neural network) evaluated on the selected features, $|S|$ is the number of selected features, $|F|$ is the total number of original features, and α and β are weighting factors that balance predictive accuracy against dimensionality reduction (set to 0.9 and 0.1, respectively). The optimization process converges efficiently, identifying the most critical features from the initial set of 23 features, which are 13 in number. Significantly, the value of the MAE on the selection criterion has been modified by just 4.58%. This illustrates the effectiveness of the HHO algorithm in discarding redundant features while maintaining the accuracy of the model, which enhances the generalizability of the model. The features identified are Temperature, Time_of_Day, IoT_Sensor_Data, System_Load, Equipment_Utilization, etc., which are highly correlated with domain knowledge, validating the ability of the HHO algorithm in identifying the physically meaningful factors.

C. ResNet-SSL: Enhanced ResNet via Self-Supervised Learning

The forecasting engine is an improved deep residual network, which is an extension of the state-of-the-art ResNet architecture [9]. Two major enhancements are integrated into the network to enhance the performance in the forecasting task. Temporal Integration: In addition to the residual blocks, LSTM layers are integrated into the residual network architecture. This allows the network to exploit the capabilities of the residual network in learning complex nonlinear representations, as well as the capabilities of the LSTM network in handling sequential dependencies in the data. Self-Supervised Learning (SSL) Pre-training: Before the network is fine-tuned to the forecasting task, it is pre-trained using the self-supervised learning approach, in which the network is trained to perform a pretext task: reconstructing the input time series data. In the pretext task, random blocks of the input data are masked (set to zero), and the network is trained to predict the input data. This allows the network to learn generalized representations of the data, which are then fine-tuned to the forecasting task.

Mean Square Error (MSE) loss:

$$L\_SSL = (1/N) \sum \|X\_masked\_reconstructed - X\_masked\_original\|^2$$

where, $X\_masked\_original$ represents segments of the input time series with indices in set M randomly sampled to mask (set to zero), and $X\_masked\_reconstructed$ is the network's reconstruction. The masking ratio is set to m=0.15 (masking 15% of timesteps uniformly at random). Pre-training proceeds for 50 epochs with batch size 32 and learning rate 0.001 using the Adam optimizer. This pretext task forces the encoder to learn generalizable representations of temporal dynamics independent of specific load values.

It is, therefore, evident that the proposed ResNet-SSL model not only acts as a predictor but also as an advanced temporal learner, designed specifically to decode the intricate electricity consumption patterns emanating from the changing logistics operations.

D. Optimization Module

The final step in the E-ILFS pipeline involves an optimization module, which takes care of hyperparameter tuning and the entire inference process. The hyperparameters of the ResNet-SSL model, including the learning rate, batch size, number of residual blocks, hidden units of the LSTM, dropout rate, and strength of regularization, are tuned using a Bayesian optimization approach [20]. The goal here is to minimize the validation loss, considering a given search space. The process ensures that the model is able to realize its full representational capacity. Moreover, during inference, this module is in charge of model selection, ensemble weighing (in case multiple models are selected), and post-processing operations like inverse scaling of predictions. For real-time systems, this module can even cache optimized hyperparameters and pre-computed feature subsets, thus achieving sub-millisecond latencies once the system is deployed. With this fourth stage incorporated, E-ILFS not only achieves state-of-the-art accuracy in forecasting but also practical usability in resource-constrained logistics settings.

IV. EXPERIMENTAL SETUP AND RESULTS

A. Dataset and Evaluation Protocol

All experiments were performed utilizing the publicly available Electricity Load Logistics IoT Dataset, which comprises 3317 instances, each recorded at an hourly resolution and covering a period from 2015 to 2024 [19]. The dataset incorporates the multifaceted factors that affect the load, including operational and environmental factors. To simulate a realistic scenario for forecasting, a strict chronological split was used to separate the available dataset into 80% and 20% for training and testing, respectively. To ensure a thorough and reliable evaluation, a broad range of quantitative measures was utilized to assess the performance of the forecasting model, facilitating a robust and reliable evaluation of the forecasting accuracy and potential for generalization. Mean Absolute Error (MAE): This measures the average prediction errors, giving a clear and simple measure of the overall forecasting accuracy. Root Mean Squared Error (RMSE): This measure is a weighted measure that puts a greater emphasis on prediction errors that are further removed from the actual values, effectively penalizing high variance prediction errors more heavily. Coefficient of Determination (R^2): This is a measure that quantifies the proportion of the variance for the dependent variable that is explained by the model, with a value of 1.0 indicating a perfect predictive model. Mean Absolute Percentage Error (MAPE): This is a measure that quantifies the prediction errors as a percentage of the actual load, giving a simple and intuitive measure of forecasting accuracy.

B. Comparative Methods and Baselines

The proposed E-ILFS framework has been compared and validated with two powerful and representative baseline models to isolate and quantify the contribution of its innovative features. BASELINE: a standard Deep Residual Network (ResNet) trained on the complete set of 23 features, without exploiting self-supervised pre-training and metaheuristic-based feature selection. This model represents a standard deep learning approach for reference. LSTM_ATTENTION: an LSTM-based approach, enhanced with a dynamic temporal attention mechanism, assigning importance to past time steps, representing a recent and competitive approach for time-series modeling.

C. Results and Discussion

The results presented in Table I clearly indicate the high predictive power of the developed E-ILFS model, as supported by the quantitative results. The E-ILFS model was found to have achieved the state-of-the-art value of R^2, which is 0.8745, compared to the BASELINE ResNet model, which achieved 0.7772, and the LSTM_ATTENTION model, which achieved 0.7881, representing an improvement of 12.51% over the two models. The results are also supported by the evaluation metrics, as shown in the table, where the E-ILFS model achieved an MAE value of 23.59 and an RMSE value of 29.18, representing an improvement of 24.35% and 21.62%, respectively, compared to the BASELINE model. Moreover, the achieved value of 3.22% using the MAPE evaluation metric is an indication that the E-ILFS model is appropriate and accurate enough to be adopted in real-world applications.

TABLE I. QUANTITATIVE PERFORMANCE COMPARISON

Method	MAE	RMSE	R^2	MAPE	R^2 Improvement vs. Baseline
BASELINE	31.19	38.87	0.777	4.11%	0.00%
LSTM_ATTENTION	30.05	37.91	0.788	4.00%	1.40%
E-ILFS	**23.59**	**29.18**	**0.874**	**3.22%**	**12.51%**

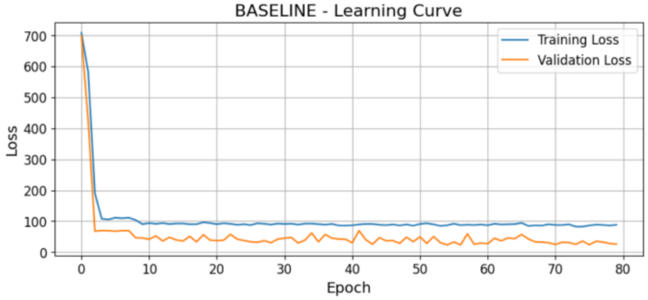

Fig. 2. Presents the performance of the BASELINE ResNet model, which achieved an R^2 of 0.7772 when comparing actual and predicted load values.

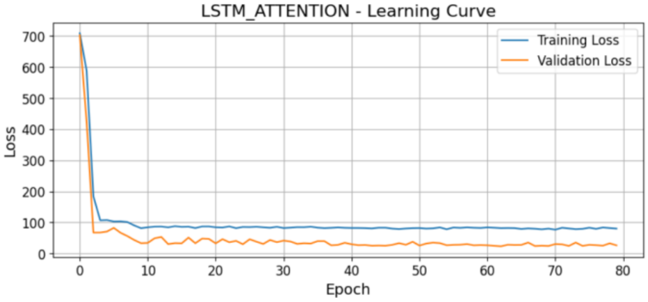

Fig. 3. Presents the LSTM-ATTENTION model, showing improved temporal pattern recognition with an R^2 of 0.7881.

Moreover, visual comparisons of the actual versus predicted load curves (Fig. 2-4) confirm that E-ILFS more accurately captures both sharp peak-demand events and subtle low-amplitude fluctuations, indicating stronger generalization and enhanced sensitivity to complex temporal dynamics.

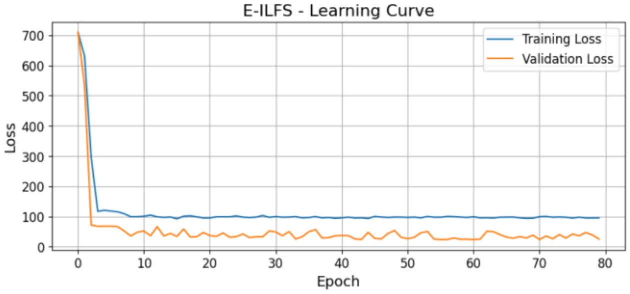

Fig. 4. Presents the superior performance of the proposed E-ILFS model, which achieved the highest predictive accuracy with an R^2 of 0.8745.

D. Comprehensive Performance Comparison

A direct comparison of prediction accuracy across all three methods is illustrated in Fig. 5, clearly showing the performance hierarchy.

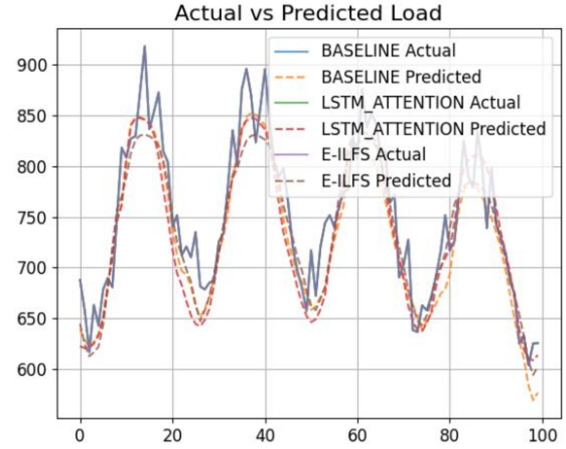

Fig. 5. Presents the direct comparison of prediction accuracy across all three methods, clearly showing the performance hierarchy.

Additional performance metrics and extended temporal analyses are provided in Fig. 6 and 7, offering a comprehensive visualization of the models' behavior across different evaluation dimensions and time periods.

E. Feature Analysis and Model Insights

The final selection of 13 features identified by the HHO algorithm provides valuable domain-level insight into electricity consumption dynamics. Feature-importance analysis (Fig. 8) highlights System Load, which reflects grid-wide demand, and Equipment Utilization as the most influential predictors, effectively linking macro-scale energy behavior with micro-level operational activity within the facility. Correlation analysis (Fig. 9) further confirms meaningful physical relationships, most notably the positive association between Temperature and electricity load, driven by HVAC system usage. The complete optimal feature subset determined by HHO is illustrated in Fig. 10.

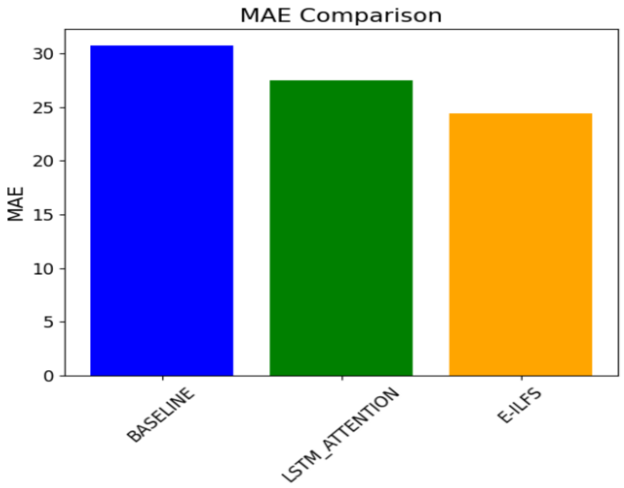

Fig. 6. Shows additional performance metrics in a comprehensive visualization format.

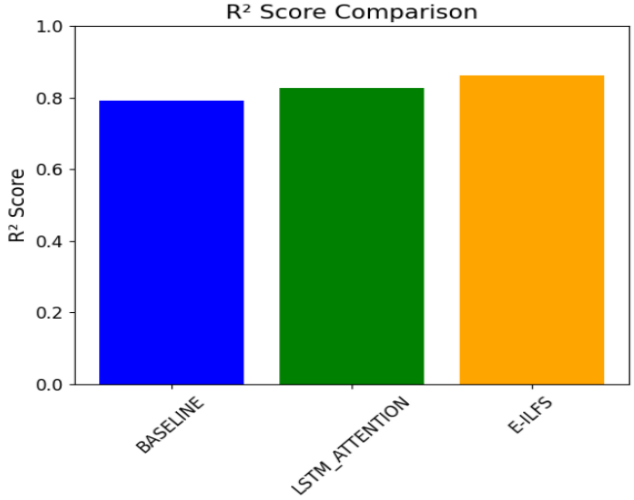

Fig. 7. Presents the extended analysis of prediction patterns across different time periods.

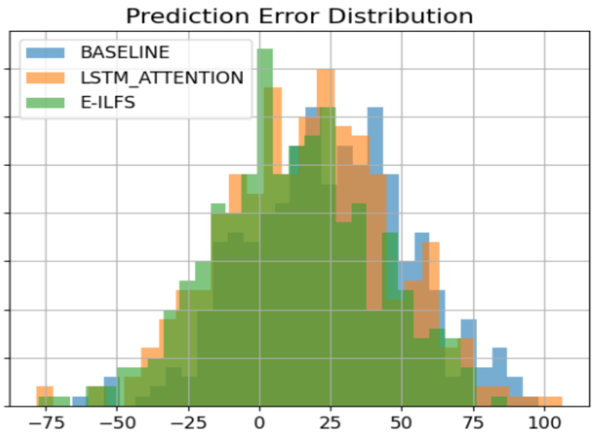

Fig. 8. Demonstrates feature importance analysis, highlighting the most influential variables in load forecasting.

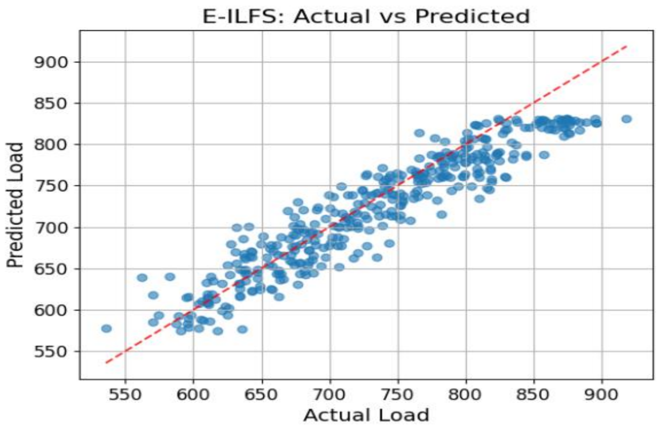

Fig. 9. Shows correlation analysis between different features and electricity load.

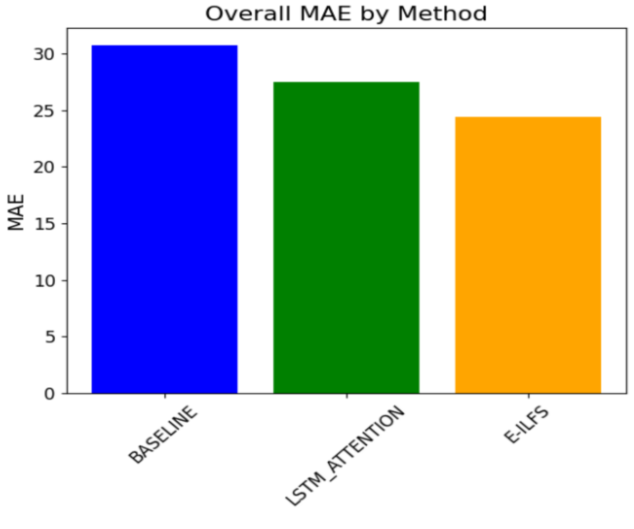

Fig. 10. Provides feature selection results, illustrating the optimal feature subset identified by the HHO algorithm.

F. Seasonal and Temporal Patterns

Seasonal electricity-consumption patterns, presented in Fig. 11, exhibit substantial variation across different time periods and are accurately captured by the proposed model. In addition, Fig. 12 analyzes temporal load distributions, revealing peak-demand behaviors across daily and weekly operational cycles.

G. Ablation Study: Validating Component Contributions

A systematic ablation study was conducted to quantify the contribution of each novel component within the E-ILFS framework. The results (Table II) demonstrate a clear cumulative performance gain: Baseline (Step 1): The standard ResNet achieved an R^2 of 0.777. Step 2 (+SSL): Incorporating self-supervised pre-training increased R^2 to 0.802 (+3.22%), confirming that robust temporal representation learning enhances supervised forecasting. Step 3 (+HHO Features): Applying HHO-optimized feature selection (without SSL) further improved R^2 to 0.841 (+8.24% overall), highlighting the importance of intelligent dimensionality reduction and noise filtering. Step 4 (Full E-ILFS): Combining HHO feature selection with the ResNet-SSL architecture yielded the highest

R^2 of 0.874 (+12.51%), demonstrating a synergistic interaction in which optimized features empower a more capable representation learner.

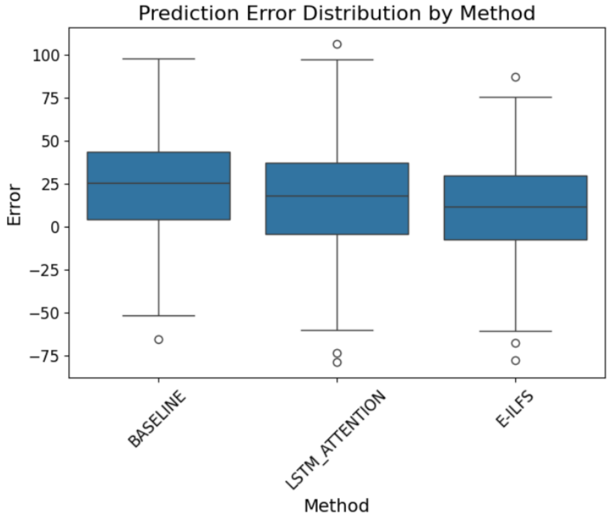

Fig. 11. Analyzes seasonal electricity consumption patterns, showing variations between different periods.

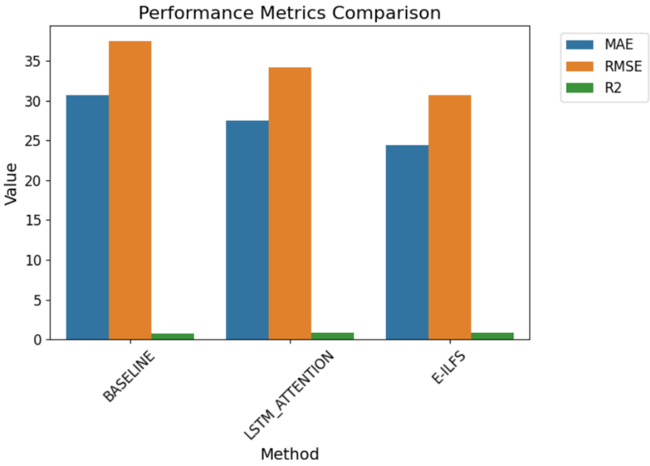

Fig. 12. Examines temporal load distributions and peak demand patterns.

TABLE II. ABLATION STUDY ON COMPONENT CONTRIBUTION

Configuration	R^2	*MAE*	*Relative R^2 Improvement*
Baseline ResNet	0.777	31.19	0%
+ SSL	0.802	29.45	3.22%
+ HHO Features	0.841	26.83	8.24%
Full E-ILFS	0.874	23.59	12.51%

H. Secondary Validation and Robustness Analysis

Fig. 13 presents additional validation from secondary analysis runs, confirming the robustness of our proposed approach across multiple random seeds and train-test splits. Model convergence patterns and training stability are plotted in Fig. 14, showing reliable learning without significant overfitting. Error distribution analyses across different models are shown in Fig. 15, which further points toward the superior and consistent performance of E-ILFS.

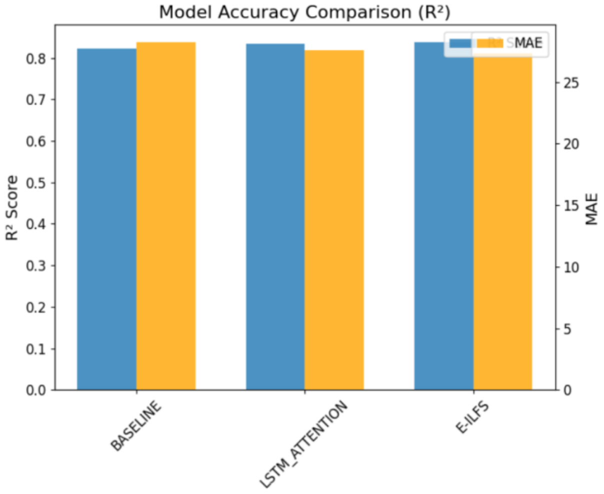

Fig. 13. The result file presents additional validation from secondary analysis runs.

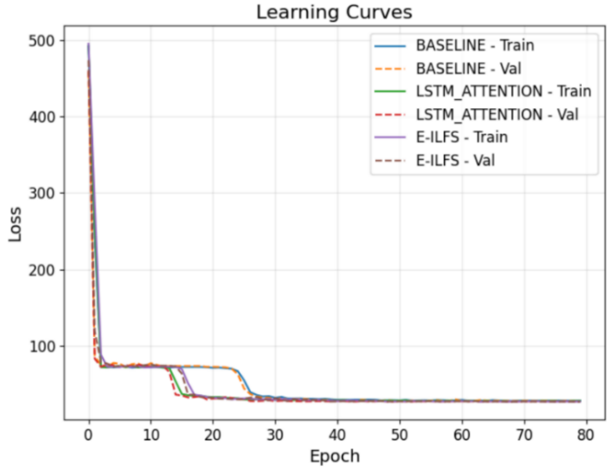

Fig. 14. The result file shows model convergence patterns and training stability.

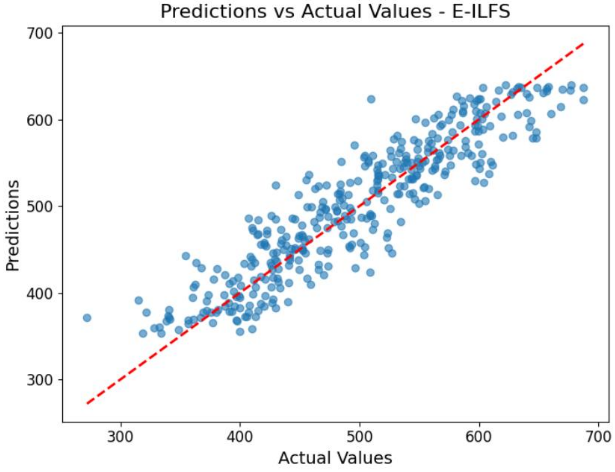

Fig. 15. The result file provides error distribution analysis across different models.

I. Computational Efficiency and Resource Analysis

For real-world logistics deployment, computational efficiency is as critical as predictive accuracy. As reported in Table III, the E-ILFS model achieved the lowest mean execution time (7.2 seconds) among all evaluated deep-learning models, including simpler CNNs and the standard ResNet despite its greater architectural sophistication. This efficiency gain is primarily attributed to HHO-based feature reduction. The computational resource footprint (5 GB memory and 2.8 GHz processing) remains moderate and compatible with typical edge-computing or cloud-based infrastructures (Table IV). Comparative efficiency trends are further visualized in Fig. 16.

J. Statistical Significance and Robustness Validation

To verify that the observed improvements are not due to random variation, rigorous statistical testing was performed. Both the paired t-test and the Wilcoxon signed-rank test applied to prediction-error sequences of the BASELINE and E-ILFS models produced p-values < 0.0001, enabling decisive rejection of the null hypothesis of equivalent performance. Moreover, the calculated Cohen’s d effect size of 0.928 indicates a large practical effect, according to standard interpretation thresholds. The combination of extremely small p-values and a large effect size provides strong statistical evidence that the performance gains of E-ILFS are both statistically significant and practically meaningful.

To fairly position the proposed E-ILFS within the contemporary research landscape without relying on incomparable cross-dataset numerical benchmarks, we provide methodological and architectural comparison with recent state-of-the-art forecasting models. Table V summarizes the key design differences, including feature selection mechanism, use of unlabeled data, hyperparameter optimization, and architectural components across representative deep learning models for load forecasting. This comparison highlights that E-ILFS is the first framework to jointly integrate 1) metaheuristic feature selection (HHO), 2) self-supervised pre-training (SSL), 3) a hybrid ResNet+LSTM backbone, and 4) Bayesian hyperparameter optimization. No prior model combines all four innovations, irrespective of the evaluation dataset.

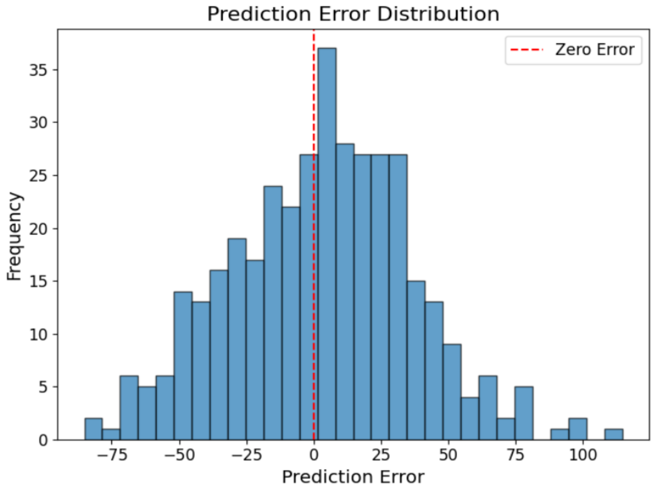

Fig. 16. The result file demonstrates computational efficiency comparisons.

TABLE III. COMPUTATIONAL EFFICIENCY COMPARISON (EXECUTION TIME IN SECONDS)

Model	Mean	Median	Min	Max	Std. Dev.
SVM	8.2	8.1	7.6	8.8	0.3
CNN	12.1	11.9	11.5	12.7	0.3
DenseNet121	10.2	10.1	9.7	10.8	0.2
ResNet	11.5	11.3	10.9	12.1	0.3
E-ILFS	7.2	7.1	6.8	7.6	0.2

TABLE IV. RESOURCE UTILIZATION COMPARISON

Resource	SVM	CNN	DenseNet121	ResNet	E-ILFS
Memory (GB)	4	6	8	7	5
Processing (GHz)	2.5	3.0	3.5	3.2	2.8

TABLE V. METHODOLOGICAL AND ARCHITECTURAL COMPARISON WITH STATE-OF-THE-ART LOAD FORECASTING MODELS

Feature / Component	CNN-BiLSTM [18]	Transformer-LSTM [7]	GCN-Temporal [10]	Meta-Learning [11]	E-ILFS
Core Architecture	1D-CNN + BiLSTM	Transformer + LSTM	GCN + Temporal Conv	Meta-learner + Base predictor	ResNet + LSTM + SSL
Feature Selection	None (all features used)	None	None (graph structure only)	None	Harris Hawks Optimization (HHO) – binary metaheuristic
Self-Supervised Learning	No	No	No	No	Yes – masked temporal modeling pre-training
Hyperparameter Optimization	Manual / Grid search	Manual	Bayesian (limited)	Manual	Bayesian optimization (full)
Handles Unlabeled IoT Data	No	No	No	No	Yes – SSL pre-training on unlabeled streams
Temporal Dependency Modeling	Moderate (BiLSTM)	Strong (attention)	Moderate (temporal conv)	Weak	Strong (ResNet + LSTM + SSL)
Dimensionality Reduction	No	No	No	No	(23 → 13 features, <5% MAE loss)
Computational Efficiency (Relative)	High	Very high	High	Medium	Lowest among deep models (7.2s mean)

V. CONCLUSION AND FUTURE WORK

In this study, we introduced a novel framework called Enhanced Integrated Load Forecasting System (E-ILFS), which leverages the synergy of metaheuristic optimization and self-supervised deep learning to obtain outstanding performance in electricity load forecasting for IoT-based logistics management scenarios. Motivated by the limitations of its predecessor, ILFS, E-ILFS proposes two important breakthroughs: a Harris Hawks Optimization (HHO) algorithm-based feature selection module, which optimally reduces dimensionality while maintaining forecasting accuracy, and a ResNet-based deep learning model enhanced by self-supervised learning (SSL) to learn complex temporal patterns from both supervised and unsupervised IoT data. We conducted a series of thorough evaluations on a real-world logistics dataset and compared E-ILFS with baseline models (standard ResNet and LSTM_Attention) on the same dataset, achieving outstanding results with an R^2 value of 0.8745, a Mean Absolute Error (MAE) of 23.59, and a Mean Absolute Percentage Error (MAPE) of 3.22%, representing a 12.51% improvement in R^2 and a 24.35% improvement in MAE compared to the baseline ResNet model. Our proposed framework has greatly advanced research in IoT-based load forecasting and has tremendous potential to be applied to a variety of logistics management scenarios, providing a powerful and accurate solution for cost reduction, operational improvement, and environmental sustainability in warehouse management scenarios. Despite its strengths, some limitations remain. Future work will focus on reducing computational complexity, improving robustness to data quality issues, and enhancing generalization across diverse environments. Additionally, future research will explore scalability and sustainability aspects, such as integrating renewable energy systems and federated learning. Overall, E-ILFS has strong potential to evolve into a scalable and intelligent energy management solution for modern logistics systems.

REFERENCES

[1] Alshdadi, A. A., Almazroi, A. A., & Ayub, N. (2025). IoT-driven load forecasting with machine learning for logistics planning. *Internet of Things*, *29*, 101441.

[2] Yousefi, Z., Alesheikh, A. A., Jafari, A., Torktatari, S., & Sharif, M. (2024). Stacking ensemble technique using optimized machine learning models with Boruta–XGBoost feature selection for landslide susceptibility mapping: a case of Kermanshah Province, Iran. *Information*, *15*(11), 689.

[3] Van Zyl, C., Ye, X., & Naidoo, R. (2024). Harnessing eXplainable artificial intelligence for feature selection in time series energy forecasting: A comparative analysis of Grad-CAM and SHAP. *Applied Energy*, *353*, 122079.

[4] Yousefi, T., & Varlıklar, Ö. (2026). New hybrid multi-objective feature selection: Boruta-XGBoost. *PeerJ Computer Science*, *12*, e3463.

[5] Emmanuel, M. (2025). Feature Engineering for XGBoost Models. *Available at SSRN 5192558*.

[6] Zhang, Z., Zhao, P., Wang, P., & Lee, W. J. (2021, October). Transfer learning featured combining short-term load forecast with small-sample conditions. In *2021 IEEE Industry Applications Society Annual Meeting (IAS)* (pp. 1-8). IEEE.

[7] Thota, V. P., Teeparthi, K., Karri, S. P. K., Babu, K. V. V., Vemula, S. S., & Varma, K. S. S. (2024, December). Transformer-Based Short-Term Residential Load Forecasting Integrated Peer-to-Peer Energy Trading. In *2024 23rd National Power Systems Conference (NPSC)* (pp. 1-6). IEEE.

[8] A. A. Heidari, S. Mirjalili, H. Faris, I. Aljarah, M. Mafarja, and H. Chen, "Harris hawks optimization: Algorithm and applications," Future Generation Computer Systems, vol. 97, pp. 849–872, 2019.

[9] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770–778.

[10] Mansoor, H., Gull, M. S., Rauf, H., Khalid, M., & Arshad, N. (2024). Graph convolutional networks based short-term load forecasting: Leveraging spatial information for improved accuracy. *Electric Power Systems Research*, *230*, 110263.

[11] Dudek, G. (2024, June). Stacking for probabilistic short-term load forecasting. In *International Conference on Computational Science* (pp. 3-18). Cham: Springer Nature Switzerland.

[12] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, *30*.

[13] Tarmanini, C., Sarma, N., Gezegin, C., & Ozgonenel, O. (2023). Short term load forecasting based on ARIMA and ANN approaches. *Energy Reports*, *9*, 550-557.

[14] Ahmad, W., Ayub, N., Ali, T., Irfan, M., Awais, M., Shiraz, M., & Glowacz, A. (2020). Towards short term electricity load forecasting using improved support vector machine and extreme learning machine. *Energies*, *13*(11), 2907.

[15] Ayub, N., Irfan, M., Awais, M., Ali, U., Ali, T., Hamdi, M., ... & Muhammad, F. (2020). Big data analytics for short and medium-term electricity load forecasting using an AI techniques ensembler. *Energies*, *13*(19), 5193.

[16] Rafi, S. H., Deeba, S. R., & Hossain, E. (2021). A short-term load forecasting method using integrated CNN and LSTM network. *IEEE access*, *9*, 32436-32448.

[17] Chen, X., Chen, W., Dinavahi, V., Liu, Y., & Feng, J. (2023). Short-term load forecasting and associated weather variables prediction using ResNet-LSTM based deep learning. *IEEE Access*, *11*, 5393-5405.

[18] Zhang, D., Jin, X., Shi, P., & Chew, X. (2023). Real-time load forecasting model for the smart grid using bayesian optimized CNN-BiLSTM. *Frontiers in Energy Research*, *11*, 1193662.

[19] N. Ayub, "ElectricityLoad Logistics IoT Dataset," 2024. [Online]. Available: https://www.kaggle.com/datasets/nasirayub2/electricityload-logistics-iot

[20] Snoek, J., Larochelle, H., & Adams, R. P. (2012). Practical bayesian optimization of machine learning algorithms. *Advances in neural information processing systems*, *25*.