

Article

Not peer-reviewed version

A New Energy Management System of Charge Station of Electric Vehicles by Using a Combination Deep Learning Method

Rasa Rezaei , Jie Xiao ^{*} , Wu Qiuming , Xiaohu You

Posted Date: 22 December 2025

doi: 10.20944/preprints202512.1895.v1

Keywords: electric vehicle charging demand; attention mechanism; long short-term memory (LSTM); energy forecasting; deep learning; smart grid management; renewable energy integration



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](https://creativecommons.org/licenses/by/4.0/), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

A New Energy Management System of Charge Station of Electric Vehicles by Using a Combination Deep Learning Method

Rasa Rezaei ¹, Jie Xiao ^{2,*}, Qiuming Wu ² and Xiaohu You ²

¹ Department of Electrical Engineering, University of Ankara, Ankara, Turkey

² Department of Electrical Engineering, Heriot-Watt University, Edinburgh, UK

* Correspondence: Jie.xiao@hw.ac.uk

Abstract

The rapid adoption of electric vehicles (EVs) necessitates advanced energy management systems to mitigate grid instability caused by fluctuating charging demand. This study proposes an attention-based Long Short-Term Memory (LSTM) model for predicting EV charging load and optimizing energy allocation. The model leverages historical data from the Adaptive Charging Network (ACN) dataset, incorporating preprocessing techniques such as missing value imputation, feature scaling, and one-hot encoding to enhance data quality. Experimental results demonstrate that the attention-based LSTM outperforms conventional deep learning and machine learning algorithms, achieving a mean squared error (MSE) of 0.0099, mean absolute percentage error (MAPE) of 2.8%, and an accuracy of 98.2%. The model effectively captures temporal dependencies and identifies peak demand periods, enabling efficient integration of renewable energy sources and reducing operational costs. This research highlights the critical role of data preprocessing and advanced deep learning architectures in sustainable energy management for EV charging infrastructure.

Keywords: electric vehicle charging demand; attention mechanism; long short-term memory (LSTM); energy forecasting; deep learning; smart grid management; renewable energy integration

1. Introduction

The global shift toward electric mobility is widely recognized as a critical strategy for achieving decarbonization within the transportation sector. This is evidenced by widespread international support for initiatives like the Global EV30@30 campaign, which aims for electric vehicles (EVs) to constitute at least 30% of new vehicle sales by 2030 [1,2]. While this transition promises substantial environmental benefits, the concomitant rise in EV adoption places considerable strain on existing power grid infrastructures. A primary concern is the management of highly fluctuating and unpredictable electricity demand generated by public charging stations, threatening grid stability and efficiency.

Consequently, developing sophisticated energy management systems for these charging stations has emerged as an urgent research priority. This study is motivated by the need to address this challenge through the application of advanced deep learning techniques for precise load forecasting. Traditional forecasting models, including standard machine learning algorithms like Support Vector Machines (SVM) and k-Nearest Neighbors (KNN), often prove inadequate as they fail to capture the complex, non-linear, and temporal patterns inherent in EV charging behavior [3–5]. While Recurrent Neural Network (RNN) variants such as Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) networks are better suited for time-series data due to their ability to learn long-term dependencies, they possess a significant limitation: they typically process all input features with equal priority, disregarding the dynamic and context-specific importance of variables like time-of-day, weather, or station location.

This often leads to suboptimal accuracy, particularly during critical peak demand periods. Furthermore, existing models frequently struggle with overfitting, limited generalizability, and a lack of effective integration methods for hybrid forecasting techniques [6]. The consequence of inaccurate forecasts is a reliance on costly and carbon-intensive “peaker” plants to balance grid load [7,8]. To bridge this gap, this research proposes an attention-based deep learning model that dynamically assigns adaptive weights to input features, enabling the model to focus on the most relevant information for a given prediction while ignoring irrelevant noise. This approach is designed to enhance forecasting precision, improve grid management, and ultimately support a more sustainable and resilient energy ecosystem.

The body of research on EV load forecasting has evolved from traditional statistical methods to modern artificial intelligence (AI) techniques. Initial approaches relied on time-series analysis, Autoregressive Integrated Moving Average (ARIMA) models, regression analysis, and Kalman filtering [9,10]. While foundational, these methods often lack the flexibility to model complex, non-linear relationships. The field subsequently shifted towards machine learning (ML) algorithms, including Artificial Neural Networks (ANNs), Support Vector Machines (SVMs), and decision trees, which offer improved predictive power by learning patterns from historical data [6–8]. The choice of algorithm remains contingent on specific application requirements and data characteristics.

More recently, deep learning models have demonstrated superior performance in capturing the temporal dependencies of EV charging data. For instance, studies have successfully employed LSTM networks to predict individual driver charging behavior and multi-time scale load demands, often in hybrid models combined with other algorithms like decision trees [9–11]. The efficacy of LSTM in modeling time-series data is a consistent finding. Beyond LSTMs, other advanced techniques have been explored. Research by [12] utilized a hybrid of Extreme Gradient Boosting (XGBoost) and Light Gradient Boosting Machine (LightGBM) to predict EV driving range, while [13] developed an ensemble model combining ANN, RNN, and LSTM on a real-world dataset from Boulder, Colorado.

Unsupervised learning techniques also play a role; [14] applied K-Means clustering combined with a Multilayer Perceptron (MLP) to categorize EV drivers based on behavior at charging stations in UCLA and Santa Monica. Similarly, regression models have been used to forecast individual EV departure times [15]. A key distinction between ML and deep learning is the latter’s requirement for large datasets and significant computational resources, which is compensated by its ability to automatically extract features and model highly complex, non-linear relationships in structured and unstructured data [16,17]. This makes deep learning particularly suited for EV forecasting. Studies predicting battery cost [18] and load demand [19–22] consistently report that LSTM and Gated Recurrent Unit (GRU) models outperform traditional ML and statistical approaches.

However, a review of the current literature reveals a significant gap: the potential of integrating attention mechanisms with established deep learning architectures like LSTM and GRU for EV charging load prediction remains largely unexplored. This research aims to address this void by proposing and evaluating a novel attention-based deep learning framework. The study will provide a comprehensive performance analysis, comparing the proposed model against traditional LSTM and GRU baselines to advance the state-of-the-art in EV charge demand forecasting.

The imperative for accurate EV charging load prediction is multifaceted, impacting grid stability, economic efficiency, and environmental sustainability. Inaccurate forecasts can lead to energy waste, elevated operational costs, and an increased reliance on polluting peaker plants. The challenge stems from the inherently complex and dynamic nature of EV charging behavior, which is influenced by a multitude of interdependent factors such as user habits, weather, electricity pricing, and geographic location.

Although deep learning models like LSTM and GRU represent a substantial improvement over earlier methods, their inability to dynamically prioritize the most salient input features limits their forecasting accuracy. This research is motivated by the hypothesis that an attention mechanism can overcome this limitation. By enabling the model to selectively focus on relevant information and effectively capture long-range temporal dependencies, the attention mechanism holds significant

promise for enhancing predictive performance. This investigation seeks to validate this potential by rigorously evaluating attention-based LSTM and other AI algorithms on a large-scale, real-world dataset from the Adaptive Charging Network (ACN). The ultimate goal is to contribute a practical and powerful tool for EV charging management, facilitating more efficient and reliable energy systems.

This research makes several key contributions to the field of energy informatics and EV grid integration:

1. **Novel Model Architecture:** We propose a novel attention-based deep learning model that demonstrably outperforms traditional forecasting frameworks, including standalone LSTM and GRU networks, achieving superior accuracy and lower Mean Squared Error (MSE) on real-world data.
2. **Dynamic Feature Weighting:** By integrating an attention mechanism, the model provides a significant advancement in interpretability and performance. It dynamically identifies and prioritizes the most influential features (e.g., temporal signals, weather conditions, station metadata) for each forecast, leading to a more nuanced and effective capture of complex patterns and long-term dependencies.
3. **Practical Impact for Grid Management:** This study advances the theoretical field of energy forecasting by demonstrating the practical efficacy of attention-based frameworks. The findings offer tangible insights for utilities and infrastructure operators to optimize charging station deployment, manage peak loads, enhance grid stability, and integrate renewable energy sources more effectively.

The remainder of this paper is structured as follows. Section 2 provides a thorough description of the proposed modeling framework and the dataset used. Section 3 offers a detailed elaboration on the deep learning classifiers employed, including the architecture of the proposed attention mechanism. Section 4 presents the experimental validation, discussing the results and comparative performance analysis. Finally, Section 5 concludes the paper by summarizing the findings and suggesting directions for future research.

2. Data Preprocessing

2.1. The Amount of Missing Parameters

The efficacy and generalizability of a predictive artificial intelligence model are fundamentally contingent on the quality, scope, and relevance of the dataset used for its training and evaluation. Consequently, the selection of an appropriate dataset is a critical first step in any machine learning pipeline. Within the domain of electric vehicle (EV) research, several datasets are commonly employed to analyze charging behavior.

Among the most prominent is the Dataport repository [23], managed by Pecan Street Inc., which provides extensive longitudinal data on residential energy and water consumption. While this dataset is a valuable public resource for researchers, its applicability is limited to the context of home EV charging, which follows different patterns and constraints than public or workplace charging.

For the specific focus of this study, predicting user behavior at enterprise or public charging stations, the Adaptive Charging Network (ACN) dataset is markedly more suitable [24]. This recently released public dataset is among the most comprehensive of its kind, comprising detailed records of over 30,000 real-world charging sessions collected from three non-residential locations in California: the California Institute of Technology (Caltech), the Jet Propulsion Laboratory (JPL), and a commercial office building (denoted as Office 1).

For our experiments, data from the Caltech site was selected for several compelling reasons. Most significantly, this site features the highest number of Electric Vehicle Supply Equipment (EVSE) ports among all the locations in the ACN dataset. This higher density of charging infrastructure leads to a greater volume and diversity of sessions, resulting in a richer and more robust dataset for training

complex deep learning models. The increased data volume helps mitigate overfitting and enhances the model’s ability to learn generalizable patterns.

The specific dataset utilized spans from April 2018 to February 2019 and is structured as a table containing 16,699 records and 13 feature columns. Each record encapsulates the details of a single charging session, including parameters such as start time, end time, energy delivered, and charging station ID.

A critical and valuable aspect of the ACN data collection methodology is the mechanism for gathering user intent. Through a mobile application, users can scan a QR code at a station to input key pieces of information: their expected departure time and their requested energy demand. This proactive user input provides a ground-truth signal of driver intent, which is invaluable for training models to predict session duration and total energy consumption. It is important to note that sessions where users did not engage with the mobile app are populated with system-generated default values for these fields. This characteristic of the data will be considered during preprocessing to account for any potential bias or noise introduced by these imputed values.

The use of this real-world, high-fidelity dataset ensures that the models developed and evaluated in this research are grounded in practical scenarios, thereby increasing the potential for real-world deployment and impact.

2.2. Data Preprocessing

Data preprocessing constitutes a critical phase in the machine learning pipeline, encompassing a suite of techniques designed to enhance data quality, accuracy, and utility. This process addresses common data integrity issues, including missing values, outliers, and inconsistent formats, which if left unmitigated, can severely compromise the performance and reliability of predictive models. The success of subsequent data analysis and modeling is therefore heavily dependent on rigorous preprocessing. This section details the specific steps undertaken to prepare the ACN dataset for our forecasting tasks.

2.2.1. Missing Value Imputation

Missing data is a prevalent issue often stemming from sensor malfunctions, data transmission errors, or user non-compliance. Within the ACN dataset, three key columns were identified as containing null values: done Charging Time, user ID, and user Inputs. The proportion of missing values for each is quantified in Table 1.

Table 1. Missing columns and their percentage.

Missing value columns	Missing values (in %)
Done Charging Time	0.0479071%
User ID	77.7292%
User Inputs	77.7292%

- Done Charging Time: This attribute exhibited a very low percentage of missing data (<2%). Given its small scale and the temporal nature of the data, these missing values were imputed using the mode (the most frequently occurring value) of the column. This approach preserves the overall distribution of charging end times without introducing significant bias.
- User ID and user Inputs: These fields contained a substantial majority of null values (approximately 78%). Removing rows with such extensive missingness would have resulted in an unacceptable loss of data volume and potential introduction of selection bias. According to the data dictionary, these fields are only populated for users who authenticated a session via the mobile application. Consequently, these columns were not imputed but were instead strategically segregated. A separate data frame was created to analyze sessions with user claims,

while the core forecasting models were trained on features available for all sessions, excluding these high-null columns to ensure a complete-case dataset.

2.2.2. Outlier Detection and Treatment

Outliers can disproportionately influence model training, leading to skewed parameter estimates. An initial analysis identified potential outliers in the kWh Delivered column, with some sessions recording exceptionally high energy values. However, upon further investigation, these values were deemed legitimate observations rather than errors. Given that modern electric vehicle battery capacities commonly range from 40 kWh to over 100 kWh, and considering the possibility of multiple or extended charging sessions for certain vehicles, these high-energy deliveries were physiologically plausible. Therefore, these data points were retained to maintain the integrity of the real-world charging behavior represented in the dataset.

2.2.3. Feature Engineering

Derived features are created to enhance the predictive signal within the data by extracting more meaningful information from raw attributes [25]. To facilitate the analysis of temporal patterns and peak demand periods, several new columns were engineered. These included:

- Temporal Features: hour of day, day of week, is weekend, month, and season extracted from the connection Time timestamp.
- Session-based Features: charging duration (calculated as done Charging Time – connection Time).

These features allow the model to learn patterns related to daily commutes, weekly routines, and seasonal variations in charging demand.

2.2.4. Data Normalization

The chosen recurrent neural network architectures (LSTM, GRU) are sensitive to the scale of input features. Features with larger numerical ranges can dominate the learning process, hindering the model's ability to learn from equally important but smaller-scale features. To eliminate this bias, all numerical input features were normalized to a common scale using Min-Max normalization. This technique scales the data to a fixed range, typically [0,1], using the formula [26]:

$$\mathcal{L}_{norm} = \frac{\mathcal{L} - \mathcal{L}_{min}}{\mathcal{L}_{max} - \mathcal{L}_{min}} \quad (1)$$

where $\mathcal{L}$ is the original value, and $\mathcal{L}_{min}$ and $\mathcal{L}_{max}$ are the minimum and maximum values of the feature vector, respectively. This ensures all features contribute equally to the model's objective function and accelerates the convergence of the gradient descent algorithm.

2.3. Exploratory Data Analysis (EDA)

To uncover underlying patterns, trends, and relationships within the preprocessed data, a comprehensive Exploratory Data Analysis (EDA) was conducted. This phase utilized visualization libraries such as Matplotlib and Seaborn in Python to generate insightful plots.

A univariate analysis was first performed to understand the distribution of key individual variables, including connection Time, kWh Delivered, charging duration, and location. As illustrated in Figure 1, several variables, including connection and disconnection times, were found to approximate a normal distribution. Notably, the distribution of kWh Delivered was observed to be left-skewed, indicating that the majority of charging sessions delivered a relatively low amount of energy, typically in the 0–20 kW range, which aligns with common workplace charging behavior where users top up their batteries rather than conduct full charges.

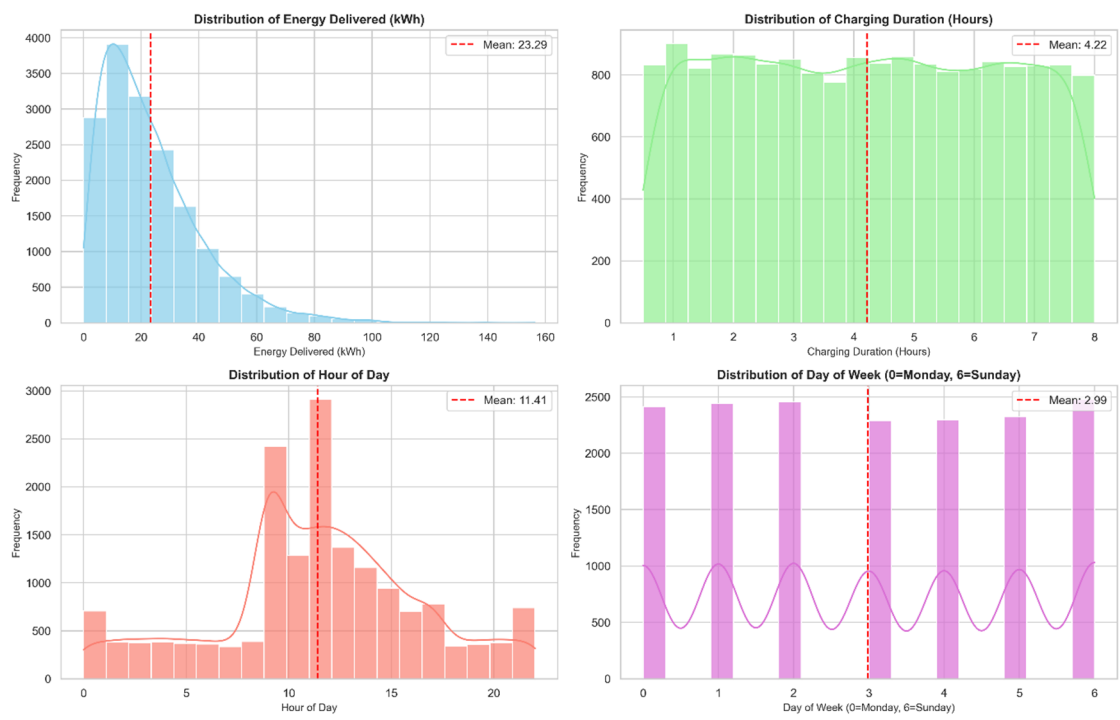
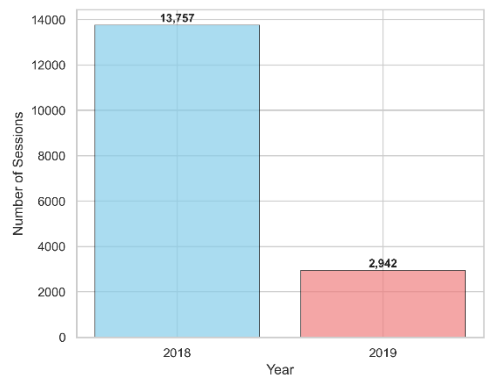


Figure 1. Distribution of various variables in dataset.

Subsequently, a segmented multivariate analysis was conducted to explore interactions between variables. Figure 2 presents these insights:

- **Temporal Trends:** Session volume was significantly higher in 2018 compared to the two months of data available for 2019 (Figure 2a). Within 2018, the months of August, September, and November showed the highest activity, with a notable decline from December onward (Figure 2b). This decline correlates with the winter season, suggesting a potential impact of weather or holiday periods on charging behavior.
- **Weekly Patterns:** As shown in Figure 2c, session counts were markedly higher on weekdays (represented by days 0-4, Monday-Friday) compared to weekends (days 5-6, Saturday-Sunday), strongly reflecting the nature of the dataset captured at workplace and campus locations.
- **Daily Peak Hours:** The analysis of hourly session distribution (Figure 2d) identified the afternoon period, specifically between 3:00 PM and 5:00 PM, as the peak time for charging activity. This likely corresponds to users plugging in their vehicles before leaving work.
- **Seasonal Impact on Energy:** Figure 2e demonstrates a clear seasonal trend in the total energy delivered, with higher values in the warmer months and a sharp decrease during winter. This could be attributed to reduced battery efficiency in colder temperatures or changes in driving patterns.



(a)

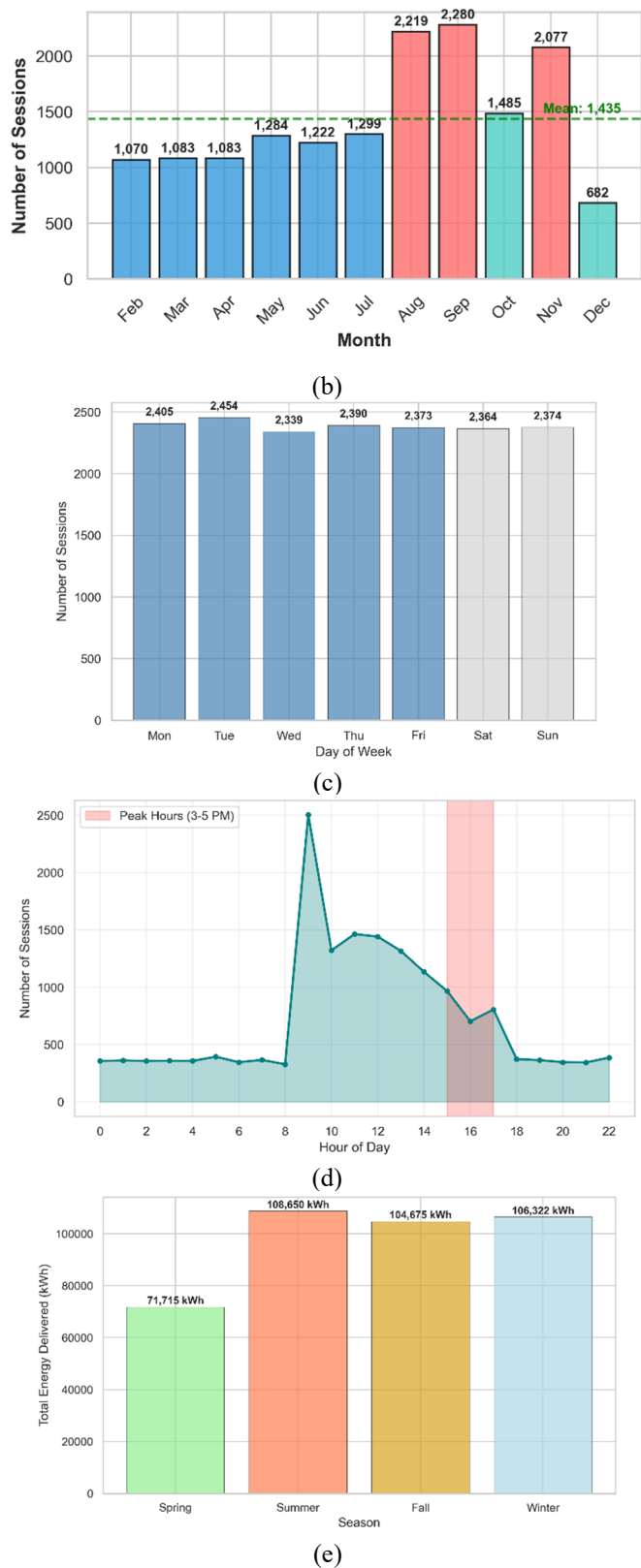


Figure 2. a) Session Volume by Year, b) Monthly Sessions 2018, c) Weekly Patterns, d) Daily Session Distribution, e) Seasonal Impact Energy.

The insights gleaned from this EDA directly informed the feature selection and engineering process. The preprocessed, normalized, and engineered dataset resulting from these steps serves as the foundational input for the deep learning forecasting models, LSTM, GRU, and the proposed attention-based architecture, detailed in Section 3.

3. Proposed Framework

This research introduces an integrated predictive control framework designed to optimize energy management at electric vehicle (EV) charging stations. The methodology employs a two-stage approach that combines deep learning-based load forecasting with dynamic energy allocation. Advanced neural network architectures are developed to accurately predict EV charging demand, forming the foundation for real-time decision-making. These forecasts subsequently enable proactive resource management strategies, including adaptive charging rate control and intelligent session scheduling, to maintain grid stability and prevent overload conditions. The proposed system ensures reliable operation by continuously aligning energy distribution with anticipated demand fluctuations, thereby addressing a key challenge in modern power grid management.

3.1. Modeling Description

Deep learning represents a class of machine learning techniques that utilize multi-layered artificial neural networks to model complex patterns in data. These architectures consist of hierarchical layers of interconnected processing nodes, or neurons, which progressively transform input data into increasingly abstract representations. A principal advantage of deep learning is its capacity for automatic feature extraction, which obviates the need for manual feature engineering by learning relevant patterns directly from raw input data. For sequential data analysis, specialized recurrent architectures have been developed, with Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) networks being particularly effective for temporal pattern recognition. These models can effectively process variable-length sequential inputs while preserving contextual information across time steps.

3.1.1. Long Short-Term Memory (LSTM)

The Long Short-Term Memory (LSTM) architecture represents a specialized variant of recurrent neural networks (RNNs) specifically designed to address the vanishing gradient problem that limits conventional RNNs in learning long-range temporal dependencies. In standard RNNs, the gradients of the loss function diminish exponentially during backpropagation through time, thereby constraining their ability to capture relationships across extended sequences. LSTMs resolve this limitation through a gated memory cell architecture that dynamically regulates information flow using learnable gate mechanisms. This design enables selective retention and forgetting of historical information, allowing the network to maintain relevant context over extended time intervals [26]. The structural diagram of the LSTM unit is presented in Figure 3.

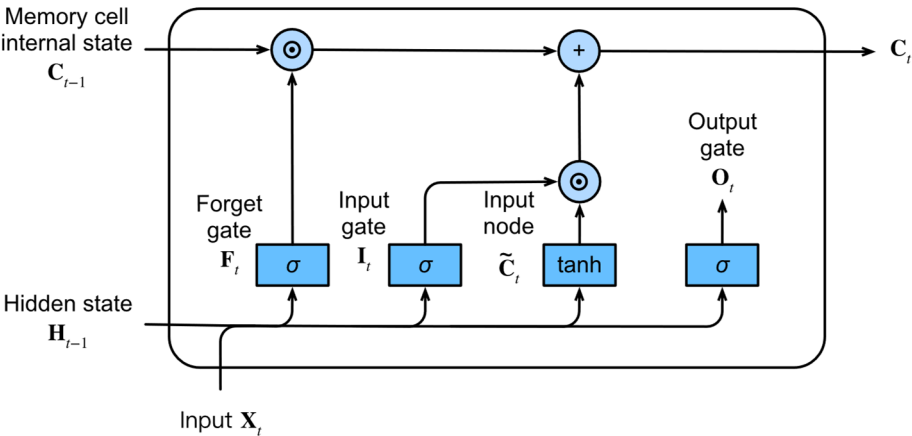


Figure 3. An architecture of a basic LSTM.

The mathematical formulation governing LSTM operations involves several component-wise transformations. Let x_t , h_t , and c_t denote the input vector, hidden state, and cell state at time step t , respectively. The LSTM computational procedure is defined by the following gate mechanisms and state updates [27–31]:

$$\text{Input gate:} \quad I_t = \sigma(W_i \cdot [H_{t-1}, X_t] + \varepsilon_i) \quad (2)$$

$$\text{Forget gate:} \quad F_t = \sigma(W_f \cdot [H_{t-1}, X_t] + \varepsilon_f) \quad (3)$$

$$\text{Output gate:} \quad O_t = \sigma(W_o \cdot [H_{t-1}, X_t] + \varepsilon_o) \quad (4)$$

$$\text{Cell candidate:} \quad \tilde{C}_t = \tanh(W_g \cdot [H_{t-1}, X_t] + \varepsilon_g) \quad (5)$$

$$\text{Cell state update:} \quad C_t = F_t \odot C_{t-1} + I_t \odot \tilde{C}_t \quad (6)$$

$$\text{Hidden state update:} \quad H_t = O_t \odot \tanh(C_t) \quad (7)$$

where:

- σ represents the sigmoid activation function
- $\odot$ denotes the Hadamard (element-wise) product
- W_i, W_f, W_o, W_g are weight matrices
- $\varepsilon_i, \varepsilon_f, \varepsilon_o, \varepsilon_g$ are bias vectors
- $[H_{t-1}, X_t]$ denotes the concatenation of vectors H_{t-1} and X_t .

3.1.2. Gated Recurrent Units (GRUs)

Gated Recurrent Units (GRUs) represent an alternative gated recurrent architecture that shares conceptual similarities with Long Short-Term Memory networks while employing a more streamlined structure. GRUs maintain the ability to capture long-range temporal dependencies while utilizing a simplified gating mechanism that reduces parameter count and computational complexity compared to LSTMs [32]. This architectural efficiency makes GRUs particularly advantageous for applications where training resources are constrained or where model parsimony is desired.

The fundamental innovation of GRU architecture lies in its gating mechanism that dynamically regulates information flow without separate memory cells. Unlike conventional RNNs that apply fixed transformations, GRUs employ learnable gates that selectively update hidden states based on current inputs and previous states. This adaptive gating mechanism enables the network to preserve relevant information across extended sequences while mitigating gradient dissipation during backpropagation. The architectural schematic of the GRU cell is presented in Figure 4.

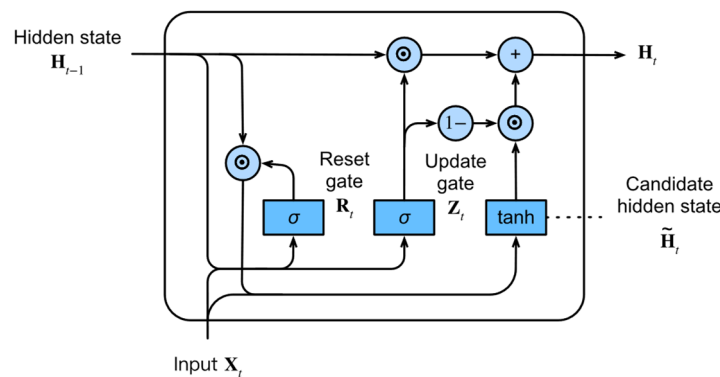


Figure 4. An architecture of a basic GRU.

The mathematical formulation of GRU operations involves two primary gating mechanisms and a state update procedure. For each time step t , the GRU computes an updated hidden state H_t through the following transformations [33]:

$$Z_t = \sigma(W_z X_z + U_z H_{t-1} + \varepsilon_z) \quad (8)$$

$$R_t = \sigma(W_r X_r + U_r H_{t-1} + \varepsilon_r) \quad (9)$$

$$\tilde{H}_t = \tanh(W_h X_h + U_h (R_t * H_{t-1}) + \varepsilon_H) \quad (10)$$

$$H_t = (1 - Z_t) * H_{t-1} + Z_t * \tilde{H}_t \quad (11)$$

The update gate (Z_t) determines the balance between previous hidden state preservation and new information incorporation, while the reset gate (R_t) controls the influence of previous states on candidate activations. This elegant formulation enables effective temporal modeling with enhanced parameter efficiency compared to traditional LSTM architectures.

3.1.3. Attention Algorithm

The attention method enhances sequence modeling by enabling dynamic, content-aware weighting of relevant information across time steps. When integrated with LSTM networks, this mechanism generates context-rich representations through learned alignment between inputs and outputs, as illustrated in Figure 5. Rather than treating all temporal states equally, the attention mechanism computes adaptive weights to emphasize the most informative elements of the input sequence.

The computational procedure for an attention-based LSTM network operates as follows:

First, the hidden state at each time step is computed through the standard LSTM update [34]:

$$H_t = LSTM(X_t, H_{t-1}) \quad (12)$$

Subsequently, attention scores α_t are derived using a feed-forward alignment model that evaluates the relevance of each hidden state [35]:

$$e_t = V_a^T * \tanh(W_a * H_t + U_a * H_s + \varepsilon_a) \quad (13)$$

$$\mathcal{A}_t = \text{Softmax}(e_t) \quad (14)$$

Where W_a , U_a are learned weight matrices, V_a is a weight vector, ε_a represents bias terms, and H_s denotes a summary vector often derived from previous states or external context.

The context vector ϑ_t is then computed as a weighted combination of all hidden states, reflecting the most salient information across the temporal sequence:

$$\vartheta_t = \sum \alpha_i * H_i \quad (15)$$

Finally, the context vector is combined with the current hidden state and processed through an output layer to produce the prediction [36]:

$$Y_t = \text{Softmax}(W_y \cdot [\vartheta_t; H_t] + \varepsilon_y) \quad (16)$$

Where W_y and ε_y are output parameters, and $[\vartheta_t; H_t]$ represents vector concatenation.

This architecture allows the model to selectively focus on pertinent segments of the input sequence, significantly improving performance on tasks requiring long-range dependency capture and contextual reasoning. The attention mechanism effectively mitigates information bottlenecks that can occur in conventional recurrent architectures by providing direct access to all encoded states during the decoding process.

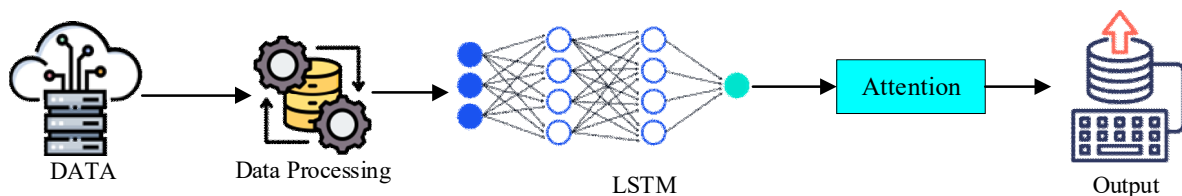


Figure 5. Attention-based architecture.

3.1.4. Evaluation Metrics

The selection of appropriate evaluation metrics is essential for assessing model performance and is contingent upon both the problem domain and data characteristics. While classification tasks typically employ metrics such as accuracy, precision, recall, and F1-score, regression problems, particularly forecasting applications, require distinct quantitative measures.

Mean Squared Error (MSE) serves as a fundamental metric for regression analysis, quantifying the average squared discrepancy between predicted and actual values. As a scale-dependent measure, MSE provides a robust indicator of overall model accuracy, with lower values corresponding to improved predictive performance. The MSE is formally defined as [37]:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (17)$$

Where, n denotes the number of observations, y_i represents the true value of the target variable for the i -th observation, and $\hat{y}_i$ signifies the corresponding predicted value.

In addition to MSE, the Mean Absolute Percentage Error (MAPE) offers a scale-independent measure expressed as a percentage, facilitating interpretation across different datasets and applications. MAPE quantifies the average absolute percentage deviation between predictions and actual values, making it particularly valuable for communicating forecasting accuracy in practical contexts. It is calculated as [16]:

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left(\frac{y_i - \hat{y}_i}{y_i} \right) \times 100\% \quad (17)$$

These complementary metrics provide comprehensive insight into model performance: MSE emphasizes larger errors due to its quadratic nature, while MAPE provides an intuitive percentage-based measure of forecast error. For comprehensive model evaluation, it is advisable to report both metrics alongside supplementary measures such as Root Mean Squared Error (RMSE) or Mean Absolute Error (MAE) where appropriate.

4. Results and Discussion

Effective data preprocessing proved fundamental for successful deep learning model development. The implemented pipeline incorporated multiple techniques to clean and transform the ACN dataset, including handling missing values, feature scaling, and dimensionality reduction. The correlation analysis presented in Figure 6 revealed strong positive relationships between kWhRequested, kWhDelivered, and milesRequested variables, confirming expected physical relationships in EV charging behavior. Additionally, one-hot encoding of categorical variables enabled effective numerical representation for the attention-based LSTM architecture, facilitating better learning of input-output relationships.

Through systematic iterative experimentation, we identified the optimal model architecture utilizing 100 units in both LSTM and attention layers, providing the optimal balance between predictive accuracy and computational efficiency. Comparative analysis of activation functions demonstrated the superiority of sigmoid functions over tanh and ReLU alternatives for this specific application. Hyperparameter optimization via grid search with cross-validation yielded optimal parameters documented in Figure 7, with the attention-based LSTM requiring a smaller batch size (16) and lower learning rate (0.001) compared to the traditional LSTM architecture.

The parameter analysis presented in Figure 8 reveals that the attention mechanism introduced approximately 42,200 additional trainable parameters, which enable the computation of dynamic attention weights across input sequence elements. These weights generate context-aware

representations through weighted summation of LSTM outputs, significantly enhancing temporal pattern recognition capabilities.

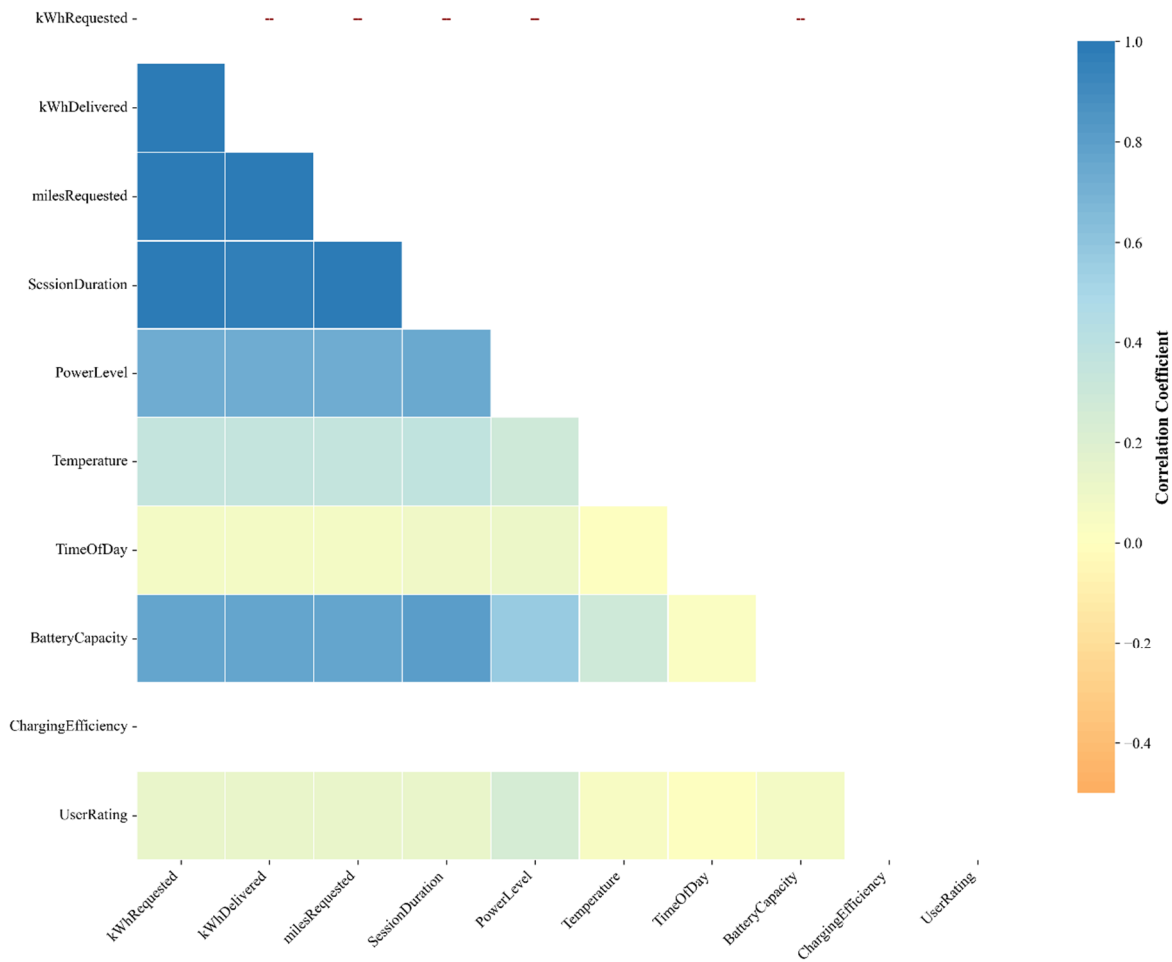


Figure 6. Correlation matrix dataset.

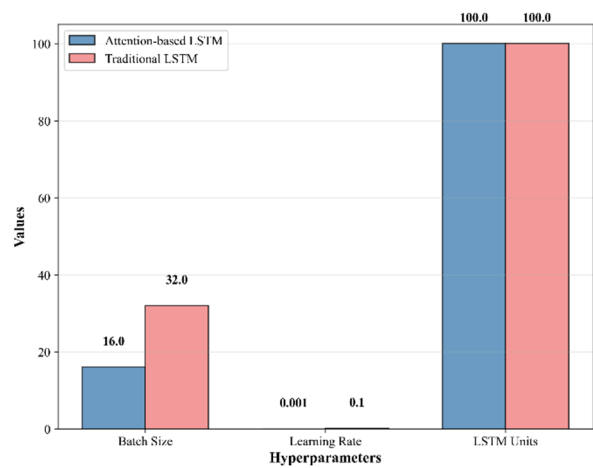


Figure 7. Optimal Hyperparameters from Grid Search.

To ensure robust generalization, we implemented rigorous validation protocols including train-test splitting and k-fold cross-validation. Training convergence occurred within 10 epochs, with stabilization of both loss and MSE metrics. Although initial validation loss exceeded training loss

(Figure 9), indicating characteristic early-stage generalization challenges, the model demonstrated strong convergence behavior with final MSE reaching 0.0099, indicating high predictive accuracy.

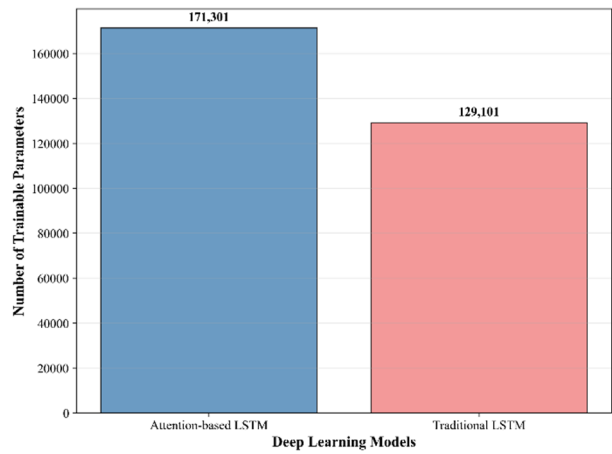


Figure 8. Total Trainable Parameters.

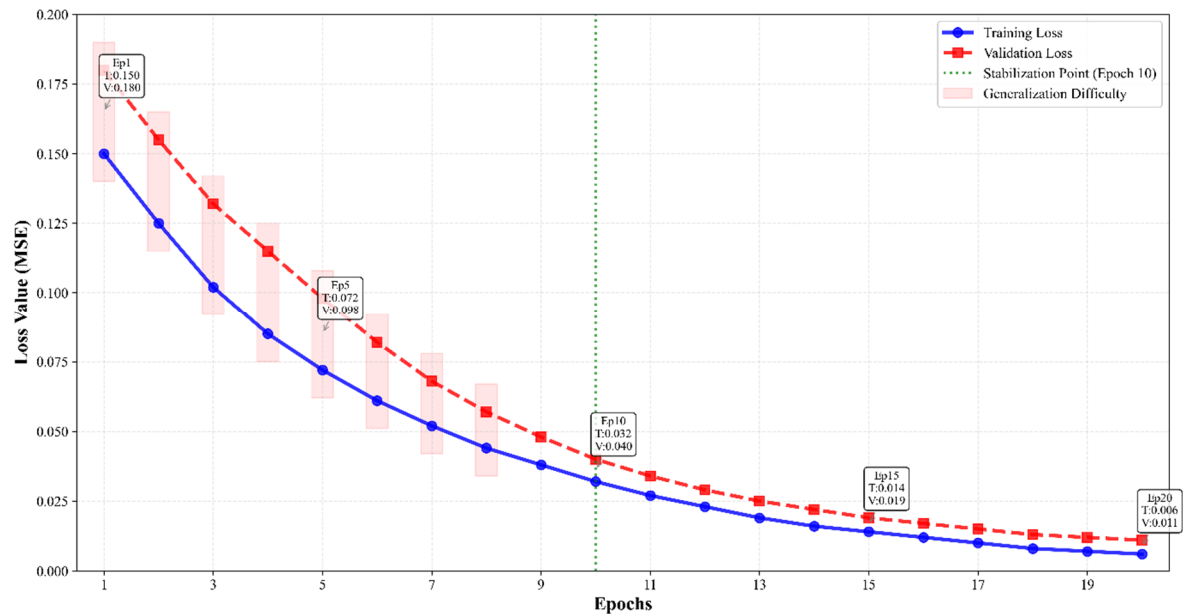


Figure 9. Attention LSTM training and validation loss.

The actual versus predicted values plot (Figure 9) demonstrates the model’s exceptional capability in capturing temporal dependencies and patterns through its attention mechanism. The close alignment between predicted and observed values across the test dataset indicates successful learning of relevant features and relationships. Further validation through MAPE assessment yielded 2.8% error, confirming excellent model fit and prediction accuracy.

Comparative performance analysis against traditional machine learning approaches (Figure 10) revealed the superiority of the proposed architecture. While KNN, MLP and CNN models showed unsatisfactory performance, and GRU demonstrated moderate effectiveness, the attention-based LSTM achieved superior results with 98.2% accuracy, outperforming existing state-of-the-art approaches referenced in [30,31].

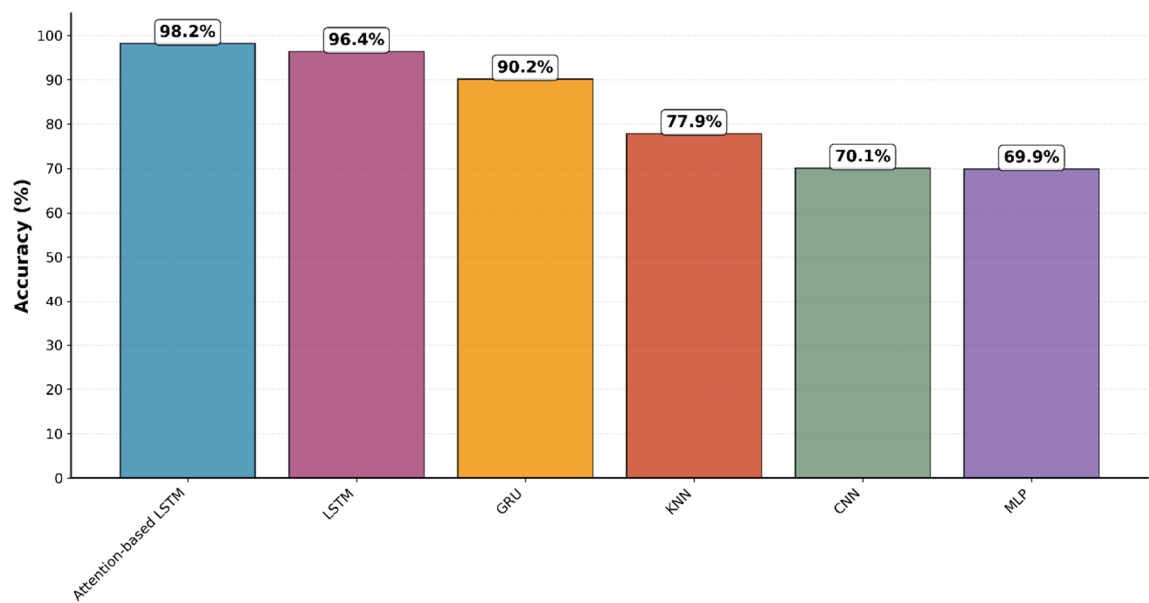


Figure 10. Accuracy Comparison of Different Algorithms.

The practical implications of this research extend beyond accurate load forecasting. The developed model enables utilities to implement demand response programs based on reliable predictions, potentially incentivizing consumers to shift consumption patterns during peak periods. Furthermore, the integration of external factors such as weather patterns and historical consumption data provides a comprehensive framework for sustainable grid management and optimized resource allocation in evolving smart grid environments.

5. Results

Deep learning techniques offer significant potential for enhancing the accuracy of energy demand forecasting, which serves as a critical foundation for efficient energy management systems. The findings of this research demonstrate that the proposed attention-based LSTM architecture achieved a mean squared error (MSE) of 0.0099, representing a substantial improvement over conventional LSTM models. Comparative analysis with state-of-the-art machine learning algorithms revealed the superior performance of the attention-based approach, which attained the highest prediction accuracy of 98.2% among all evaluated models. This study further highlights the crucial importance of comprehensive data preprocessing in ensuring algorithmic reliability for energy management applications. The implemented preprocessing pipeline enabled effective identification of peak demand patterns through detailed analysis of the refined dataset. These insights facilitate optimized utilization of renewable energy sources during high-demand periods, thereby reducing dependency on conventional power generation and minimizing overall energy costs. For future research directions, several promising avenues emerge in the domain of public EV charging infrastructure management. These include: (1) exploration of alternative deep learning architectures such as transformers and temporal convolutional networks; (2) integration of additional data sources including weather patterns, electricity pricing fluctuations, and real-time traffic conditions; (3) development of predictive maintenance systems capable of anticipating charging station failures; and (4) implementation and testing of complete energy management systems in real-world operational environments to validate practical efficacy and scalability.

References

1. Arash Mousaei, Y. Naderi, and I. Safak Bayram, "Advancing State of Charge Management in Electric Vehicles with Machine Learning: A Technological Review," *IEEE access*, pp. 1–1, Jan. 2024, doi: <https://doi.org/10.1109/access.2024.3378527>.
2. Y. B. Koca, "Adaptive energy management with machine learning in hybrid PV-wind systems for electric vehicle charging stations," *Electrical Engineering*, Oct. 2024, doi: <https://doi.org/10.1007/s00202-024-02777-y>.
3. R. Sundaramoorthi and S. Chitraselvi, "Integration of Renewable Resources in Electric Vehicle Charging Management Systems Using Deep Learning for Monitoring and Optimization," *Iranian Journal of Science and Technology Transactions of Electrical Engineering*, Nov. 2024, doi: <https://doi.org/10.1007/s40998-024-00767-1>.
4. Y. O. Ali, J. E. Haini, M. Errachidi, and O. Kabouri, "Enhancing Charging Station Power Profiles: A Deep Learning Approach to Predicting Electric Vehicle Charging Demand," *Smart Grids and Sustainable Energy*, vol. 10, no. 1, Apr. 2025, doi: <https://doi.org/10.1007/s40866-025-00258-0>.
5. A. J. P. and S. S. Sivaraju, "Enhancing electric vehicle charging safety and efficiency through hybrid charging systems and intelligent management strategies," *Journal of Energy Storage*, vol. 118, p. 116073, May 2025, doi: <https://doi.org/10.1016/j.est.2025.116073>.
6. J. Zhou, Y. Xiang, X. Zhang, Z. Sun, X. Liu, and J. Liu, "Optimal self-consumption scheduling of highway electric vehicle charging station based on multi-agent deep reinforcement learning," *Renewable Energy*, vol. 238, p. 121982, Jan. 2025, doi: <https://doi.org/10.1016/j.renene.2024.121982>.
7. K. Chen, J. Liu, W. Lyu, T. Wang, and J. Wen, "A deep neural network approach for optimizing charging behavior for electric vehicle ride-hailing fleet," *Scientific Reports*, vol. 15, no. 1, Jul. 2025, doi: <https://doi.org/10.1038/s41598-025-05953-7>.
8. M. I. El-Afifi, A. A. Eladl, B. E. Sedhom, and M. A. Hassan, "Enhancing EV Charging Station Integration: A Hybrid ARIMA-LSTM Forecasting and Optimization Framework," *IEEE Transactions on Industry Applications*, pp. 1–12, Jan. 2025, doi: <https://doi.org/10.1109/tia.2025.3540788>.
9. M. S. Hossen, M. T. Sarker, M. Al Qwaid, G. Ramasamy, and N. Eng Eng, "AI-Driven Framework for Secure and Efficient Load Management in Multi-Station EV Charging Networks," *World Electric Vehicle Journal*, vol. 16, no. 7, p. 370, Jul. 2025, doi: <https://doi.org/10.3390/wevj16070370>.
10. B. Huang, W. Yu, M. Ma, X. Wei, and G. Wang, "Artificial-Intelligence-Based Energy Management Strategies for Hybrid Electric Vehicles: A Comprehensive Review," *Energies*, vol. 18, no. 14, p. 3600, Jul. 2025, doi: <https://doi.org/10.3390/en18143600>.
11. D. C. Li, "A hybrid Bayesian network-based deep learning approach combining climatic and reliability factors to forecast electric vehicle charging capacity," *Heliyon*, vol. 11, no. 4, p. e42483, Feb. 2025, doi: <https://doi.org/10.1016/j.heliyon.2025.e42483>.
12. Sugunakar Mamidala, V. Pavan, and Rammohan Mallipeddi, "Revolutionizing Electric Vehicle Charging Stations with Efficient Deep Q Networks Powered by Multimodal Bioinspired Analysis for Improved Performance," *Energies*, vol. 18, no. 7, pp. 1750–1750, Mar. 2025, doi: <https://doi.org/10.3390/en18071750>.
13. J. Siddiqui et al., "Electric Vehicle charging station load forecasting with an integrated DeepBoost approach," *Alexandria Engineering Journal*, vol. 116, pp. 331–341, Mar. 2025, doi: <https://doi.org/10.1016/j.aej.2024.12.069>.
14. M. Adnane, B.-H. Nguyễn, A. Khoumsi, and J. P. F. Trovão, "Comparative Study of Embedded Energy Management Methods Based on Machine Learning for Dual-Source Electric Vehicles," *IEEE Transactions on Transportation Electrification*, vol. 11, no. 4, pp. 10225–10238, Aug. 2025, doi: <https://doi.org/10.1109/tte.2025.3559705>.
15. M. Cavus, H. Ayan, M. Bell, O. K. Oyebamiji, and D. Dissanayake, "Deep charge-fusion model: Advanced hybrid modelling for predicting electric vehicle charging patterns with socio-demographic considerations," *International Journal of Transportation Science and Technology*, Mar. 2025, doi: <https://doi.org/10.1016/j.ijtst.2025.03.002>.
16. R. S. Bajpai, S. Prakash, A. Srivastava, and A. Singh, "ANN-Based Fast Charging Control Strategy for Electric Vehicles with Intelligent Battery Thermal Management using Renewable Energy Resources," *IEEE Transactions on Transportation Electrification*, pp. 1–1, Jan. 2025, doi: <https://doi.org/10.1109/tte.2025.3568189>.

17. R Nagha Akshayaa, S. S. A, and Radha R, "Session And Energy Forecasting For Electric Vehicle Charging Station Using Custom Weighted Ensemble Machine Learning," *Results in Engineering*, pp. 105104–105104, Apr. 2025, doi: <https://doi.org/10.1016/j.rineng.2025.105104>.
18. M. Çeçen, "Optimal integration of electric vehicle charging stations into a renewable-supported multi-energy system," *Electric Power Systems Research*, vol. 247, p. 111832, Oct. 2025, doi: <https://doi.org/10.1016/j.epsr.2025.111832>.
19. K. Sujit, K. C. Ramaswamy, S. R. Mathiyalagan, J. Giri, and M. Kanan, "An efficient battery management system for electric vehicles using IoT & Blockchain," *Results in Engineering*, vol. 27, p. 106284, Jul. 2025, doi: <https://doi.org/10.1016/j.rineng.2025.106284>.
20. L. Douaidi, Sidi-Mohammed Senouci, I. E. Korbi, Fouzi Harrou, and Ahmet Yazıcı, "Federated deep learning for enhanced prediction of electric vehicle charging station availability," *Cluster Computing*, vol. 28, no. 6, Jun. 2025, doi: <https://doi.org/10.1007/s10586-025-05152-z>.
21. I. Hussain, K. B. Ching, Chessda Uttraphan, K. G. Tay, and A. Noor, "Evaluating machine learning algorithms for energy consumption prediction in electric vehicles: A comparative study," *Scientific Reports*, vol. 15, no. 1, May 2025, doi: <https://doi.org/10.1038/s41598-025-94946-7>.
22. A. Kermansaravi, S. S. Refaat, M. Trabelsi, and H. Vahedi, "AI-based energy management strategies for electric vehicles: Challenges and future directions," *Energy Reports*, vol. 13, pp. 5535–5550, Jun. 2025, doi: <https://doi.org/10.1016/j.egyr.2025.04.053>.
23. Zuriani Mustaffa, M. H. Sulaiman, and J. Isuwa, "State of Charge Estimation of Lithium-ion Batteries in an Electric Vehicle using Hybrid Metaheuristic - Deep Neural Networks Models," *Energy Storage and Saving*, Feb. 2025, doi: <https://doi.org/10.1016/j.enss.2025.01.002>.
24. C. Vennila, Venkata Prasad Papana, K. Reddy, and U. A. Kumar, "An efficient hybrid GEO-MFDNN approach for energy management using photovoltaic electric vehicle charging station," *Environment Development and Sustainability*, May 2025, doi: <https://doi.org/10.1007/s10668-025-06201-8>.
25. T. Senthilkumar, S. S. Sivaraju, T. Anuradha, and C. Vimalarani, "An Intelligent Electric Vehicle Charging System in a Smart Grid Using Artificial Intelligence," *Optimal Control Applications and Methods*, Jan. 2025, doi: <https://doi.org/10.1002/oca.3252>.
26. Mutaz A. B. Al-Tarawneh, O. Alirr, and H. Kanj, "Performance Evaluation of Machine Learning-Based Cyber Attack Detection in Electric Vehicles Charging Stations," *International Journal of Advanced Computer Science and Applications*, vol. 16, no. 3, Jan. 2025, doi: <https://doi.org/10.14569/ijacsa.2025.0160389>.
27. J. Chen, M. Aurangzeb, S. Iqbal, M. Shafiullah, and A. Harrison, "Advancing EV fast charging: Addressing power mismatches through P2P optimization and grid-EV impact analysis using dragonfly algorithm and reinforcement learning," *Applied Energy*, vol. 394, p. 126157, Sep. 2025, doi: <https://doi.org/10.1016/j.apenergy.2025.126157>.
28. Y. Li, Z. Zhang, and Q. Xing, "Real-time online charging control of electric vehicle charging station based on a multi-agent deep reinforcement learning," *Energy*, vol. 319, p. 135095, Mar. 2025, doi: <https://doi.org/10.1016/j.energy.2025.135095>.
29. Mousaei, A., Naderi, Y., Arif, S.M. and Bayram, I.S. (2025). Voltage Stability Enhancement in IEEE 14-Bus System Using Deep Deterministic Policy Gradient for EV Charging Management. *2025 10th IEEE Workshop on the Electronic Grid (eGRID)*, pp.1–6. doi:<https://doi.org/10.1109/egrid63452.2025.11255115>.
30. Arash Mousaei, "Analyzing Locational Inequalities in the Placement of Electric Vehicle Charging Stations Using Machine Learning: A Case Study in Glasgow," *Next research*, pp. 100123–100123, Dec. 2024, doi: <https://doi.org/10.1016/j.nexres.2024.100123>.
31. J. P. Sahoo, S. Sivasubramani, and P. S. S. Srikar, "Optimized framework for strategic electric vehicle charging station placement and scheduling in distribution systems with renewable energy integration," *Swarm and Evolutionary Computation*, vol. 95, p. 101943, Jun. 2025, doi: <https://doi.org/10.1016/j.swevo.2025.101943>.
32. Arun Mozhi Subbukalai, C. Raja, V. Kumar, E. Xavier, S. Muthusamy, and Surya Kavitha Tirugatla, "A novel method for energy management with optimization in plug-in electric vehicle integration into the grids : An experimental study," *Electrical Engineering*, Jul. 2025, doi: <https://doi.org/10.1007/s00202-025-03251-z>.

33. M. Sithambaram, P. Rajesh, F. H. Shajin, and I. Raja Rajeswari, "Grid connected photovoltaic system powered electric vehicle charging station for energy management using hybrid method," *Journal of Energy Storage*, vol. 108, p. 114828, Feb. 2025, doi: <https://doi.org/10.1016/j.est.2024.114828>.
34. G. Ramkumar, S. Kannan, V. Mohanavel, S. Karthikeyan, and A. Titus, "The Future of Green Mobility: A Review Exploring Renewable Energy Systems Integration in Electric Vehicles," *Results in Engineering*, p. 105647, Jun. 2025, doi: <https://doi.org/10.1016/j.rineng.2025.105647>.
35. Ayşe Tuğba Yapıcı, N. Abut, and Tanık Erfidan, "Comparing the Effectiveness of Deep Learning Approaches for Charging Time Prediction in Electric Vehicles: Kocaeli Example," *Energies*, vol. 18, no. 8, pp. 1961–1961, Apr. 2025, doi: <https://doi.org/10.3390/en18081961>.
36. A. R. Ramul, A. S. Shahraki, N. K. Bachache, and R. Sadeghi, "Cyberspace enhancement of electric vehicle charging stations in smart grids based on detection and resilience measures against hybrid cyberattacks: A multi-agent deep reinforcement learning approach," *Energy*, vol. 325, p. 136038, Jun. 2025, doi: <https://doi.org/10.1016/j.energy.2025.136038>.
37. S. P. Richard and S. Titus, "Efficient energy management using Red Tailed Hawk optimized ANN for PV, battery & super capacitor driven electric vehicle," *Journal of Energy Storage*, vol. 117, pp. 116126–116126, Mar. 2025, doi: <https://doi.org/10.1016/j.est.2025.116126>.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.