ISSN: 2688-8645

A Review of the Application of Machine Learning in the Prediction of Electric Vehicle Charging Demand

Yixin Wang *

Manchester Metropolitan Joint Institute, Hubei University, Wuhan, Hubei, 430062, China

* Corresponding author Email: wangyixin20040408@qq.com

Abstract

This paper systematically reviews the current application status, technical paths, and future trends of machine learning in electric vehicle (EV) charging demand forecasting. Firstly, starting from the background of global energy transformation and EV popularization, it points out the crucial role of accurate charging demand forecasting in grid stability and facility optimization. Secondly, it summarizes the research progress of machine learning technology in three major scenarios: intelligent management of charging piles, battery state evaluation, and charging load forecasting. It focuses on analyzing the advantages and limitations of models such as random forest, LSTM, GCN, and Transformer in capturing spatiotemporal features and improving prediction accuracy. Furthermore, it discusses core challenges such as data quality, model generalization, real-time performance, and multi-source data fusion, and proposes future research directions such as lightweight model design, transfer learning, federated learning, and V2G collaborative optimization. Finally, through multi-model comparison and technical route analysis, it provides theoretical support and practical reference for building an integrated "vehicle-pile-grid" smart energy system.

Keywords

Electric Vehicles (EVs); Charging Demand Forecasting; Machine Learning; Spatiotemporal Modeling; Data-driven; Smart Grid.

1. Introduction

1.1. Background Information

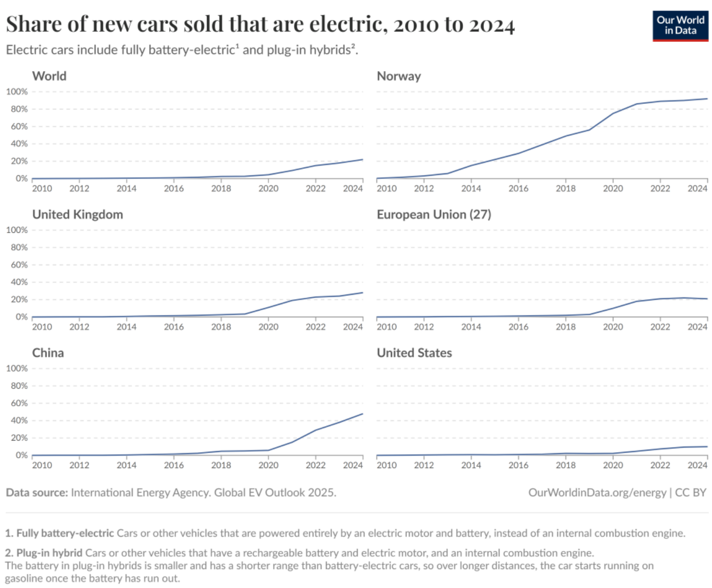

Figure 1. Proportion of new energy power in newly sold cars [1]

The energy crisis and environmental pollution issues are becoming increasingly severe. Fuel-powered vehicles emit harmful substances such as carbon monoxide and nitrogen oxides through their exhaust, and produce greenhouse gas carbon dioxide, exacerbating climate warming. Therefore, the accelerated development of clean energy vehicles is being promoted, and electric vehicles, as an important representative, are rapidly replacing traditional fuel-powered vehicles.

In 2023, the global electric vehicle (EV) stock reached 40,066,000 units, accounting for a significant portion of the newly sold vehicles that year, as shown in Figure 1[1]. The EV stock in China's mainland reached 21,800,800 units, accounting for 54.41% of the global EV stock[2]. In China, by the end of 2023, the national new energy vehicle (NEV) stock had reached 30,410,000 units, accounting for 6.07% of the total vehicle fleet[3].

1.2. Typical Application Scenarios and Cases

This research provides key technical support for the sustainable development of the electric vehicle industry, facilitating the deep integration of charging demand forecasting with smart grids, enabling bidirectional interaction between electric vehicles and the grid, optimizing grid stability and renewable energy consumption capacity, and constructing an integrated smart energy system of "vehicle-charging pile-grid"[2].

2. The Core Technology of Machine Learning in the Prediction of EV Charging Demand

2.1. Real-time Data Detection and Analysis of Charging Piles and Batteries

Through monitoring and analyzing real-time data from charging piles, fault diagnosis, condition assessment, and intelligent scheduling can be achieved, thereby enhancing the operational efficiency of the charging network.

2.1.1. Sensor Detection and Wireless Communication Transmission

By installing sensors on electric vehicles, parameters such as voltage, current, temperature, and charging and discharging status are collected in real-time, and the data is transmitted to cloud servers using wireless communication technology, enabling real-time data monitoring of charging piles and electric vehicles[4]. On the server side, big data technology is employed to integrate multi-source heterogeneous data, achieving precise monitoring and analysis of the operating status of charging piles.

2.1.2. Data Processing and Feature Extraction

Due to the characteristics of noise and nonlinearity often present in the raw data collected by sensors, preprocessing, filtering, and feature extraction are required. Preprocessing involves steps such as denoising and filtering to eliminate interference components in the data [4-6].

In the preprocessing stage, methods such as median filtering (a nonlinear smoothing technique that sets the gray value of each pixel to the median of the gray values of all pixels within a certain neighborhood window around that point), mean filtering (a typical linear filtering algorithm that involves assigning a template to the target pixel in the image, where the template includes its surrounding neighboring pixels, and then replacing the original pixel value with the average value of all pixels in the template), and wavelet denoising (which eliminates noise through short waves) can be employed to remove random noise and systematic noise from the data.

The comparison of different preprocessing processes is shown in Table 1.

ISSN: 2688-8645

Table 1. Comparison of three filtering and denoising methods

Algorithm	Principle	Advantages	Disadvantages
Median Filter	Replace pixel value with the median of the neighborhood	Strong noise resistance, edge preservation	Weak smoothing
Mean Filter	Weighted average of neighborhood pixel values	Remove Gaussian noise, simple	Blur edges
Wavelet Denoising	Threshold processing after wavelet transform	Multi - scale analysis, detail preservation	Complex computation

A common feature extraction method is principal component analysis, which can be selected and adjusted according to the characteristics of the data and the type of fault, in order to achieve the best feature extraction effect.

2.1.3. Application of Machine Learning Algorithms

The most widely adopted algorithm among them is the Random Forest algorithm[4,7,8]. As an industrial-grade statistical machine learning algorithm, Random Forest models complex nonlinear relationships by integrating multiple decision trees. It exhibits excellent generalization ability and predictive performance in classification and regression tasks across industries such as industry, healthcare, and finance.

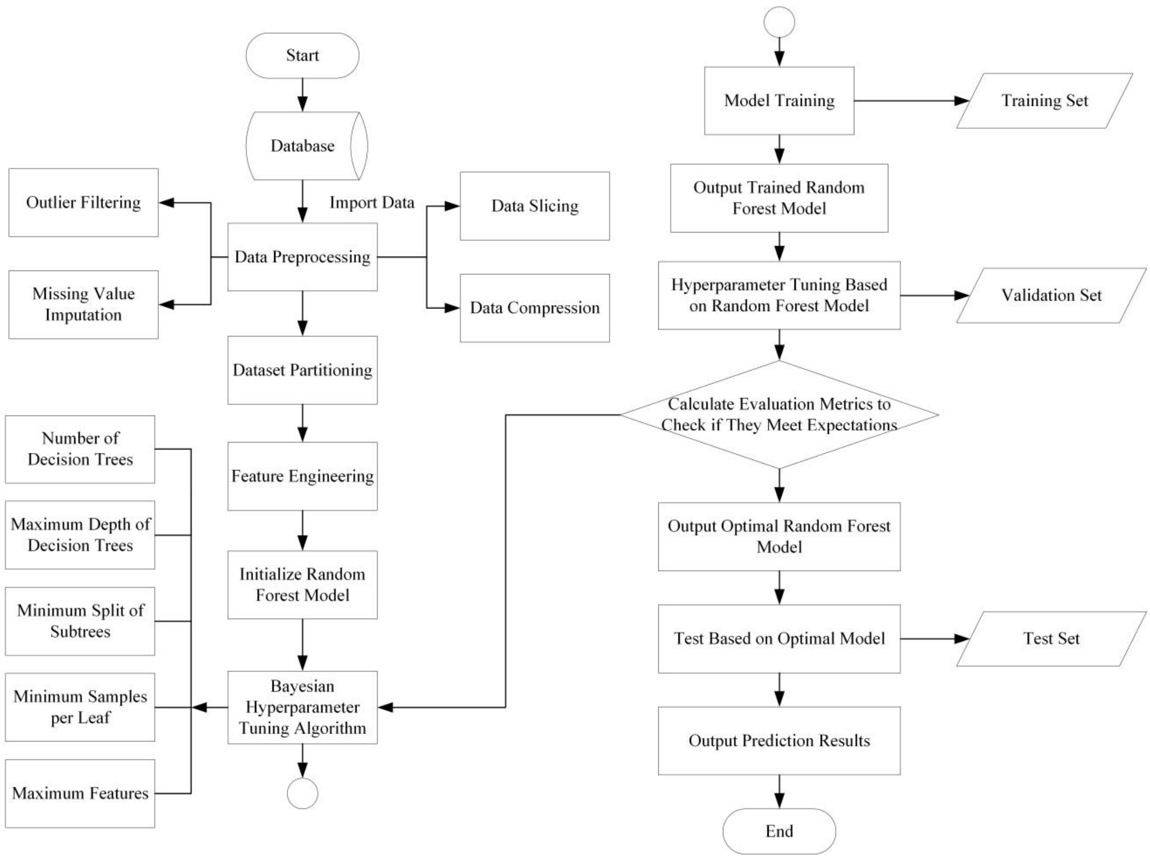

Figure 2. Modeling SOC process of electric vehicles based on the random forest algorithm

Its core mechanism comprises three parts: First, it employs the Bootstrap method to extract a fixed number of samples (allowing duplicates) from the original training set, generating

ISSN: 2688-8645

independent subsets. Each subset contains approximately 63.2% unique samples to introduce randomness into the sample space. Second, during the splitting of decision tree nodes, it randomly selects features based on the square root of the total number of features (for example, selecting $\sqrt{M}$ features when the total number of features is M). The optimal splitting feature is selected using the Gini coefficient (for classification trees) or squared error (for regression trees), limiting the feature search space to create feature space perturbations. Third, the process of repeated sampling and feature selection generates a large number of decision trees. For classification tasks, the results of multiple trees are integrated using majority voting, while for regression tasks, arithmetic averaging is used to enhance model robustness. The randomness of the algorithm is specifically reflected in two aspects: sample space perturbations (generating differentiated training subsets through Bootstrap sampling) and feature space perturbations (limiting the candidate feature set to enhance the diversity of base learners).

For regression problems, the Classification And Regression Tree (CART) is commonly used as the base learner, with splitting features and thresholds selected based on the criterion of minimizing squared error. The modeling steps are as follows: specify CART regression tree as the base learner and configure parameters such as the number of trees, depth, and number of candidate features. Perform bootstrap sampling with replacement on the training set to generate Bootstrap subsets, randomly select features to construct a single CART regression tree, and repeat until the condition of the number of trees is met. Finally, the random forest model is obtained by averaging the outputs of all regression trees, achieving high-precision estimation of continuous variables such as battery state of charge (SOC).

2.2. Analysis of User Charging Behavior

In the field of charging load forecasting, researchers predict charging loads in different time periods and regions by modeling user charging behavior and combining time series and spatial correlation features, providing a basis for grid dispatching and facility planning [9-11].

2.2.1. Cluster Analysis

The clustering analysis method in machine learning has significant advantages in the field of discovering user charging behavior patterns [7,8,12]. By mining the inherent structure and similarity of data, it automatically divides users with similar charging behavior characteristics into different groups, thereby revealing the diversity of charging patterns and user preferences.

By integrating multi-dimensional behavioral data such as user charging time, location, and power, cluster analysis can effectively identify users' inherent charging patterns, providing a reference for personalized charging services. The bivariate mixture Gaussian model clustering method constructs a mixture Gaussian probability model to simultaneously analyze the variables of user's starting and ending charging times, achieving a refined depiction of charging behavior and confirming the diversity of user behaviors. By introducing the l-p norm as a measure of dissimilarity between matrices, the reliability of clustering results is enhanced by quantifying the degree of difference in charging behavior matrices among different users[12]. This achieves automated mining of charging patterns and lays the foundation for behavioral pattern analysis.

Compared to traditional statistical models, clustering analysis places greater emphasis on the inherent structure of data, exhibiting stronger flexibility and adaptability. By integrating diverse data, innovative algorithms, and measurement methods, it deepens our understanding of charging patterns and provides data support for formulating personalized service strategies. In the future, further optimizing the applicability of clustering methods in large-scale high-dimensional data scenarios is expected to drive new breakthroughs in user charging behavior analysis.

ISSN: 2688-8645

2.2.2. Construction of User Mental Model

In the field of electric vehicles, the application of user psychological models is extensive and significant. Early research employed mechanistic models, statistical models, and cluster analysis to analyze users' behavioral characteristics such as charging time, location, and frequency, and to explore the underlying psychological motivations, such as range anxiety and charging preferences[7,13]. These models reflect users' psychological expectations and decision-making patterns under different travel needs and battery states. Meanwhile, in the demand response mechanism, psychological factors such as users' awareness and sensitivity to peak and off-peak electricity prices, as well as their preferences for charging time selection, are also taken into account to construct a psychological response model for users towards incentive measures, guiding users to adjust their charging time.

With the development of distributed renewable energy, users have transformed into prosumers with power generation capabilities, reflecting their psychological demands for energy self-sufficiency, cost reduction, and improved energy utilization efficiency. In electricity market transactions, users' considerations of market electricity prices, transaction costs, and energy supply stability also influence their decision-making behavior. Furthermore, users' psychological factors such as acceptance, trust, and recognition of virtual power plants and aggregators affect their willingness to participate in the electricity market and their interactive relationships.

The fuzzy psychological model is used to analyze users' charging decisions. This model consists of an input layer, a fuzzy processing layer, and an output layer, incorporating Z-type function, bilateral Gaussian function, and S-type function as membership functions, respectively represented as follows:

$$f_z(x; z_1, z_2) = \begin{cases} 1 & x \le z_1 \\ 1 - 2\left(\frac{x - z_1}{z_2 - z_1}\right)^2 & z_1 \le x \le \frac{z_1 + z_2}{2} \\ 2\left(\frac{x - z_1}{z_2 - z_1}\right)^2 & \frac{z_1 + z_2}{2} \le x \le z_2 \\ 0 & x \ge z_2 \end{cases} \tag{1}$$

$$f_{g2}(x; \mu, \sigma) = \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \tag{2}$$

$$f_s(x; s_1, s_2) = \begin{cases} 0 & x \le s_1 \\ 2\left(\frac{x - s_1}{s_2 - s_1}\right)^2 & s_1 \le x \le \frac{s_1 + s_2}{2} \\ 1 - 2\left(\frac{x - s_1}{s_2 - s_1}\right)^2 & \frac{s_1 + s_2}{2} \le x \le s_2 \\ 1 & x \ge s_2 \end{cases} \tag{3}$$

In the formula, z_l and z_2 are the relevant parameters in the Z-shaped function; μ and σ are the parameters in the bilateral Gaussian function; s_I and s_2 are the relevant parameters in the S-shaped function.

In addition, by analyzing factors such as SOC_{n+1}, time-of-use electricity pricing, and parking duration, we can determine charging behavior and the choice between fast and slow charging. The measurement of range anxiety primarily relies on whether the remaining battery capacity at the end point is sufficient for the next journey consumption, as shown in the following formula:

ISSN: 2688-8645

$$SOC_{n+1} = \frac{SOC_n * E - E_v * S_{n \to n+1}}{E} \tag{4}$$

In the formula, SOC_{n+1} represents the predicted state of charge at the end of the next route; SOC_n represents the remaining battery capacity at the current destination; E represents the initial battery capacity of the EV; E_v represents the EV consumption per kilometer, with the unit of kW·h/km; $S_{n \to n+1}$ represents the mileage of the next route.

Meanwhile, considering factors such as user driving experience, travel OD attributes, and route complexity, we analyze users' dependence on in-vehicle navigation and the psychological basis for driving route selection, as well as the impact of psychological factors such as charging anxiety, satisfaction with sufficient charging duration during functional area stays, and other factors on charging decisions. These studies contribute to optimizing the layout of charging facilities, formulating reasonable charging strategies, and enhancing user experience.

2.3. Various Models for Capturing Charged Charges

Models such as Random Forest, LSTM, GCN, and Transformer proposed by scholars both domestically and internationally have improved prediction accuracy by capturing the spatiotemporal characteristics of charging load[4,7,8,10,14].

2.3.1. Spatiotemporal Graph Convolutional Network Model [10]

The data on charging demand not only exhibits complex time dependencies but is also significantly influenced by the spatial dependencies among different charging stations. Graph Convolutional Networks (GCN) excel at capturing spatial relationships within graph structures, while Long Short-Term Memory (LSTM) has notable advantages in time series prediction. A hybrid model combining the two can be utilized to extract the spatio-temporal characteristics between charging stations. Feature construction and data preprocessing integrate the geographical location and time series information of charging stations as inputs. Subsequently, the geospatial data is gridded and transformed into a graph structure, where GCN is employed to extract spatial relationship features and capture dependencies between adjacent charging stations. Following this, LSTM processes time series data to identify the time dependencies of charging demand. Finally, the model outputs prediction results, which are evaluated using error metrics such as Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE). The representation of the two-layer GCN model is as follows:

$$f(A, X) = \sigma(\hat{A} ReLU(\hat{A} X W_0) W_1) \tag{5}$$

In the formula, X represents the feature matrix, which contains the features of each node in the graph; A represents the adjacency matrix, indicating the connectivity between nodes in the graph; $\widehat{A} = \widehat{D}^{-\frac{1}{2}} \widehat{A} \widehat{D}^{-\frac{1}{2}}$ represents the preprocessing step; D is a 1-degree matrix, $\widehat{D} = \Sigma \widehat{A}_{ij}$ used to normalize the adjacency matrix A to avoid the excessive influence of node degree (i.e., the number of neighbors) on the convolution result; $\widehat{A} = A + I_N$ is a matrix with self-connection structure, where I_N is the identity matrix, representing the connection between each node and itself (self-loop); W_0 and W_1 represent the weight matrices of the first and second layers, σ(·), Relu( ) is the activation function.

Although the model has achieved good results in predicting the charging demand of electric vehicles, its performance may be affected by the quality and completeness of data. Future research could explore how to utilize a broader range of data sources for training.

2.3.2. Construction of Transformer LSTM Model [10]

Due to its inherent complexity and volatility, electric vehicle load forecasting poses higher demands on model processing and modeling capabilities, especially in scenarios with multi-source input data and long time series. Traditional forecasting models struggle to accurately capture global information and long-term dependencies in complex dynamics.

To enhance prediction accuracy, a TransLSTM model based on improved long short-term memory (LSTM) is proposed, integrating the strengths of Transformer and LSTM. This model captures global contextual information and models multi-dimensional input feature relationships through the self-attention mechanism of Transformer, integrating them into the hidden state representation. Leveraging the temporal modeling capabilities of LSTM, it further analyzes the hidden state, deeply explores the global dependencies of sequential data, and thereby obtains high-precision prediction results.

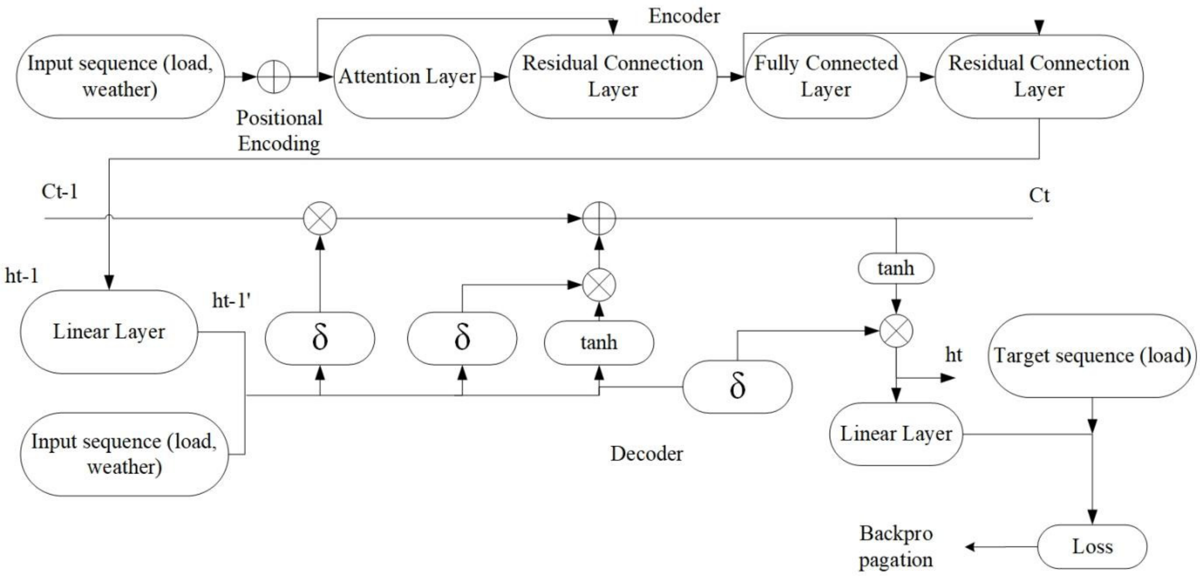

Figure 3. TransLSTM structure

2.4. The Data-driven Role of Mining Hidden Information

The data-driven approach involves cleaning and processing vast amounts of operational data (such as vehicle speed, battery state of charge, etc.), to extract implicit information to enrich model inputs[4,5].

2.4.1. Interpolation Smoothing

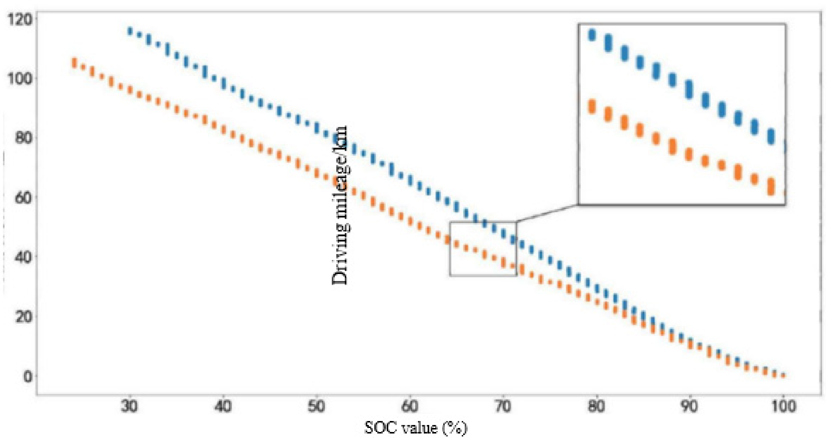

Figure 4. Before SOC data interpolation

Interpolation is defined as the process of interpolating a continuous function based on discrete data, such that the continuous curve passes through all given discrete data points. If, during a certain journey, the resolution of the original SOC data is 1%, it cannot reflect SOC changes smaller than 1% [8]. This can result in the same SOC value continuously corresponding to different cumulative driving distances within a single driving segment. For example, Figure 3 shows the relationship between cumulative mileage and SOC for two driving segments of a certain vehicle. It can be observed that because the SOC data changes in units of 1%, discrete vertical lines appear in the graph, which is not conducive to subsequent modeling.

Therefore, we consider interpolating SOC based on battery power (the product of battery voltage and battery current).

In a unit time, it is approximately assumed that the SOC decrease rate is proportional to the battery power. For each driving segment, take the data with SOC value c (0<c<100%) in the driving segment, let the number of rows of the taken data be n, the power battery voltage and power battery current data be sets U, I, $\hat{c}_j$ respectively, and SOC value after interpolation processing be j, where j is an integer and 0<j≤n. Then the SOC data interpolation formula is:

$$\hat{c}_i = c + \frac{\sum_{k=1}^{j} U_k I_k}{\sum_{k=1}^{n} U_k I_k} \tag{6}$$

After interpolation processing, the curve becomes smoother.

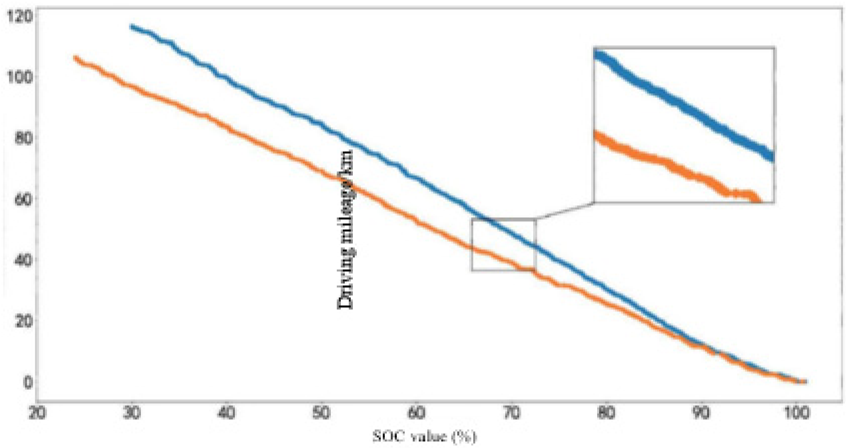

Figure 5. After SOC data interpolation

2.4.2. Crawler Retrieves POI

Web crawling technology can obtain POI data from Gaode Map, providing rich geospatial data support for research[9]. For example, researchers have utilized Python web crawling combined with the Gaode Map API to crawl POI data from Shaoxing City on Gaode Map, and classified these data into four functional zones: residential area, working area, commercial area, and public service area. This process involves filtering, extracting, and classifying map data, enabling researchers to construct models that reflect the actual functional distribution of the city based on the characteristics of different functional zones.

Through the acquisition and analysis of POI data, researchers can better simulate the travel patterns and charging needs of electric vehicle users, providing a scientific basis for subsequent planning and layout of charging facilities. This method has become an indispensable tool in the field of data-driven urban research and transportation planning. It not only improves the efficiency of data collection but also ensures the accuracy and timeliness of data, providing a solid data foundation for research on electric vehicle charging demand prediction.

2.5. Multi-model Comparison

In terms of temporal processing, LSTM excels at capturing long-sequence dependencies thanks to its gating mechanism, making it suitable for tasks such as short-term charging load forecasting[4,7,15]. However, it faces limitations such as the long-sequence gradient problem and weak global information. Bi-LSTM enhances the accuracy of temporal modeling by incorporating both past and future information through bidirectional modeling, albeit with a large number of parameters and slow training speed[7,9]. Transformer, based on the self-

ISSN: 2688-8645

attention mechanism, can efficiently process multi-dimensional inputs and capture global feature correlations, often combined with LSTM for multi-dimensional feature load forecasting[7,10]. However, it has weak local dynamic capture capabilities and high computational resource requirements. TransLSTM combines the advantages of Transformer and LSTM, taking into account both global correlations and temporal dependencies, making it suitable for multi-source input or long-term complex charging load forecasting[8,10]. However, it faces challenges such as complex structure and the risk of overfitting.

For graph-structured data, GCN-LSTM combines graph convolution and LSTM to jointly model spatial correlation and temporal features, which can be used for charging demand prediction with spatial correlation[8,9]. However, graph structure preprocessing is complex, and its efficiency is low on large-scale graphs. GCN focuses on extracting spatial correlation features from graph-structured data, also relying on the graph structure, and its efficiency is not high when dealing with large-scale graphs[8,9].

In clustering and classification algorithms, K-means, as a distance-based clustering algorithm, is simple and efficient, and can be used for preliminary classification of user behavior patterns and division of charging scenarios [7,8]. However, it requires a preset number of clusters and is sensitive to initial values. Cluster analysis can automatically discover data similarity, facilitating user behavior mining and charging scenario classification, but it demands high requirements for algorithm and parameter selection [7,8]. Support Vector Machine (SVM) excels in small-sample nonlinear classification tasks such as battery fault diagnosis, with strong high-dimensional generalization ability. However, parameter tuning is complex and computationally expensive[15].

In reinforcement learning-related models, multi-agent deep reinforcement learning possesses complex system interaction decision-making capabilities, suitable for multi-agent collaborative power distribution system optimization, but its convergence is affected by environmental non-stationarity and there is a limit on the number of agents; reinforcement learning algorithms (such as DQN) can learn strategies in dynamic environments, enabling real-time decision-making and adaptive dynamic scenarios, but training is unstable and data demand is high[13]. Markov Decision Process (MDP), as a reinforcement learning decision-making framework, can be used for modeling the charging scheduling decision-making process, but it suffers from the curse of dimensionality in large-scale problems and relies on environmental models[13].

Integrated learning models each have their unique characteristics. Random Forest exhibits high-dimensional robustness and resistance to overfitting, making it suitable for feature selection and preliminary modeling in load forecasting[4,7,8]. However, it suffers from low efficiency and weak time series processing when dealing with large-scale data. Gradient Boosting Decision Tree (GBDT), as an integrated model of gradient boosting decision trees, excels in handling nonlinear data with high accuracy and precision, making it applicable for high-precision regression tasks such as SOC estimation and mileage prediction [4,8]. However, it faces the risk of overfitting and is sensitive to parameters. XGBoost achieves efficient and accurate regression and classification of large-scale data through regularized gradient boosting, supports custom loss functions, but its parameters are complex and it tends to overfit[4,6]. LightGBM, leveraging histogram and feature parallelism techniques, efficiently processes large-scale data with low memory consumption, making it suitable for tasks such as large-scale mileage prediction[8]. However, its performance is weak on small datasets and it is sensitive to parameters. CatBoost excels in handling categorical features and has high generalization ability, directly processing classification and regression tasks with categorical features[8]. However, its training speed is slow and it relies on parameters. As a multi-granularity cascade forest integration framework, gcForest achieves high accuracy without requiring complex parameter tuning, making it suitable for tasks such as SOC estimation on small-scale data[4]. However, its structure is complex and training time is long.

Furthermore, deep learning models excel in complex feature extraction tasks such as fault diagnosis and load forecasting due to their automatic nonlinear feature extraction capabilities, albeit with high demands on data and computational resources[7,15]. YOLOv5, as a real-time object detection convolutional network, can achieve high-speed and high-precision detection, making it suitable for real-time detection of foreign objects in wireless charging [6]. However, it performs weakly in detecting small and occluded targets and has a large model size.

3. Realistic Dilemmas and Technological Research Directions

Despite significant progress in technology application, core challenges persist. In terms of data quality, electric vehicle charging data often faces issues such as high noise and numerous missing values, making efficient data preprocessing and cleaning techniques fundamental challenges for model training[5]. The model's generalization ability is insufficient, as existing models are mostly trained based on data from specific regions or vehicle models, making it difficult to adapt to different regional traffic patterns, climatic conditions, and vehicle model differences. The contradiction between real-time requirements and computational efficiency is prominent, with complex deep learning models having low deployment efficiency at the edge, making it difficult to meet the real-time dispatching needs of the power grid [13]. The complexity of multi-source data fusion is significant, and charging demand is influenced by multiple factors such as traffic flow, meteorological conditions, and user behavior. How to effectively integrate heterogeneous data (such as geographic information, real-time traffic conditions, weather data, etc.) still requires breakthroughs.

Looking ahead, research in this field can focus on five major directions: First, optimizing the structure and algorithms of machine learning models to enhance prediction accuracy and computational efficiency through lightweight design [14]. Second, deepening multi-source data fusion to explore the spatiotemporal coupling relationship between meteorological, traffic, user behavior data, and charging demand[7]. Third, developing real-time prediction technology to achieve dynamic updating of charging load through edge computing and distributed architecture[13]. Fourth, enhancing model generalization ability by integrating multi-source data using transfer learning and federated learning to enhance cross-scenario adaptability[8]. Fifth, promoting the deep integration of charging demand prediction with smart grids, enabling bidirectional interaction between electric vehicles and the grid through V2G technology, optimizing grid stability and renewable energy consumption capacity, and building an integrated smart energy system of "vehicle-charging pile-grid". Breakthroughs in these directions will provide key technical support for the sustainable development of the electric vehicle industry[2].

References

[1] Statista. (2025). Electric vehicles worldwide - Statistics & facts. https://www. statista. com/ topics/ 1010/electric-mobility/#statisticChapter.

[2] National Development and Reform Commission. (2023). Building a high-quality charging infrastructure system to support the high-quality development of new energy vehicles - Expert interpretation of the Guiding Opinions on Further Building a High-Quality Charging Infrastructure System. http://www.ndrc.gov.cn/xxgk/jd/jd/202306/t20230629_1358020.html.

[3] Ministry of Public Security of the People's Republic of China. (2025). The number of new energy vehicles reaches 31.4 million. https://www.gov.cn/ zhengce/jiedu/ tujie/202501/ content_ 6999527. htm.

[4] Gan, Z. H. (2022). Research on SOC estimation method of pure electric vehicle based on machine learning [Master's thesis, South China University of Technology].

[5] Li, Y. T. (2024). Research on electric vehicle charging load prediction based on big data. Automobile Forum, 155, 155–157.

[6] Qian, Q., Chen, H., Zheng, Y., et al. (2023). Foreign object detection method for wireless charging of electric vehicles based on machine learning. Modern Electronics Technique, 46(13), 43–48.

[7] Hou, L. N., Zhang, X., Luo, X. F., et al. (2024). A review of data-driven modeling and optimization for electric vehicle charging scheduling. Electric Engineering, 2024(21).

[8] Su, J. S., Li, Y. H., Qin, H., & Fan, J. K. (2020). Electric vehicle driving range prediction based on machine learning model. In *Proceedings of 2020 SAE - China Congress* (pp. 685–689).

[9] Gao, L. X., Gao, Y. P., Zhang, L., et al. (2024). Dynamic charging demand prediction of electric vehicles integrating spatio-temporal big data and user psychological analysis. In 2024 (2nd) Urban Power Grid Technology Innovation Conference.

[10] Geng, P., Yang, H. J., Shi, Z. X., et al. (2024). Electric vehicle charging demand prediction based on spatio-temporal graph convolutional network. Journal of Traffic Engineering, 24(11), 37–45.

[11] Zhao, W. B., Fan, X. D., Sun, S. B., et al. (2025). Scenario-based analysis and ultra-short-term prediction method for electric vehicle charging load. Electric Power Information and Communication Technology, 23(2), 28–37.

[12] Straka, M., Piatriková, L., van Bokhoven, P., et al. (2021). A matrix approach to detect temporal behavioral patterns at electric vehicle charging stations. Transportation Research Procedia, 55, 1353–1360.

[13] Wang, R. J., Gao, H. J., Luo, L. B., et al. (2025). A review of optimal operation research of new distribution systems based on deep reinforcement learning. Electric Power Automation Equipment. DOI:10.16081/j.epae.202505008

[14] Wei, S., & He, H. L. (2024). Research on fault diagnosis method of electric vehicle power battery under big data background. Automotive Safety Technology, (4), 74–76.

[15] Yang, L. (2023). Research on intelligent charging detection and automatic control system design for pure electric vehicles. Automobile Design, (11), 110–112.