World Electric Vehicle Journal

MDPI

Article

Design Optimization of Home Electric Vehicle Chargers Based on User Review Mining and Explainable Machine Learning

Yao Zhao [1], Yujia Pan [1], Jue Wang [1], Zekun Lu [2], Yulin Wang [1], Shunhe Chen [1,*] and Kaida Chen [1,*]

[1] College of Landscape Architecture and Art, Fujian Agriculture and Forestry University, 15 Shangxiadian Road, Cangshan District, Fuzhou 350002, China
[2] School of Architecture and Planning, Chongqing University, Chongqing 400045, China
* Correspondence: 000q151015@fafu.edu.cn (S.C.); kaidachenwhu@gmail.com (K.C.)

Abstract

As electric vehicles become widespread, home EV chargers have emerged as a key interface between household energy use and daily mobility. However, their design optimization remains insufficiently informed by large-scale user feedback. This study develops a review-driven, interpretable machine learning framework to identify design priorities for home EV chargers. Of the 26,763 reviews collected from the JD e-commerce platform, 23,893 were retained after cleaning. BERTopic extracted raw topics, which were consolidated into ten design dimensions through independent coding, inter-coder agreement assessment, and consensus adjudication. A structured large language model protocol then transformed the reviews into evidence-constrained, aspect-level semantic proxy variables representing evaluative direction and intensity. Coding reliability was evaluated against dual-coder annotations, while a matched absence-as-zero specification examined sensitivity to the treatment of unmentioned dimensions. Platform ratings were subsequently introduced as the prediction target, and repeated data partitions and cross-model SHAP comparisons were used to assess partition- and model-level stability. Charging Performance, Operational Stability, Perceived Product Quality, and Operational Convenience and Portability consistently ranked as the most important factors associated with platform-rated satisfaction. In contrast, Installation Friendliness and After-sales Service showed asymmetric attribution patterns characterized by stronger low-value penalties than high-value gains. The framework supports translating online review evidence into product-level design priorities, while emphasizing that SHAP identifies predictive associations rather than causal effects.

Keywords: residential charging; online reviews; electric vehicle infrastructure; large language model; explainable machine learning

Check for updates

Academic Editor: Michael Fowler

Received: 18 May 2026
Revised: 12 July 2026
Accepted: 28 July 2026
Published: 30 July 2026

Copyright: © 2026 by the authors. Published by MDPI on behalf of the World Electric Vehicle Association. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

1.1. Research Background

With the continuing advancement of the global energy transition and transport decarbonization, electric vehicles have evolved from policy-driven alternatives into widely adopted mobility options [1]. According to the International Energy Agency's Global EV Outlook 2025, global electric vehicle sales exceeded 17 million units in 2024, accounting for more than 20% of new car sales worldwide, with China contributing over 11 million units [1]. These figures indicate that electric vehicles have entered a stage characterized by both rapid expansion and broad market diffusion. At this stage, vehicle performance alone

does not fully characterize user evaluations; charging stability, convenience, and overall experience are also closely associated with satisfaction [2].

Existing studies on electric vehicle charging facilities have mainly focused on charging types, policy mechanisms, market structures [3], and infrastructure expansion, typically treating charging facilities as functional nodes within energy and transport systems. Fabianek and Madlener noted that users evaluate charging experiences across multiple dimensions, including ease of use, functionality, comfort, price transparency, and availability [4]. Studies based on real residential charging data further show that home charging exhibits relatively stable daily patterns in start time, charging duration, and energy consumption intensity, which differ markedly from those in public charging contexts [5]. Hardman and Tal found that discontinuation of plug-in electric vehicle use was significantly associated with the lack of convenient home charging conditions and dissatisfaction with charging convenience [6]. These findings suggest that home EV chargers are not merely scaled-down versions of public charging facilities, but independent product systems closely coupled with residential space, daily routines, and subsequent maintenance. Although methods such as KANO and AHP have been used to investigate user requirements and functional optimization for public EV charging stations, design optimization research on residential EV chargers remains limited, particularly analyses based on authentic user feedback, such as online reviews [7].

The distinct user experience characteristics of residential charging indicate the need to identify design dimensions from user feedback. For products used at scale and over extended periods, user reviews provide an important entry point for identifying real design problems [8]. Yang et al. argued that online reviews contain valuable information about user experience and relational structures that can support product design and strategic decision-making [8]. Qin et al. demonstrated that user feedback can be translated into evidence for design improvement through review-driven pathways for product redesign and innovation [9]. An et al. used BERTopic and deep clustering networks to extract consumer intentions, related products, and their advantages and disadvantages from reviews, providing a technical approach for deriving insights from reviews [10]. In an automotive case, Liu et al. combined BERTopic with sentiment analysis to identify user requirements, indicating that review data can support more fine-grained requirement mining [11]. However, data-driven design often lacks a sufficient link between data analysis and final design decision-making, making the translation of data evidence into design knowledge a persistent challenge [12]. Related studies have also shown that deep learning has been applied in product design practice to support analytical tasks and improve user experience [13].

Extracting structured evidence for design optimization from reviews differs from general sentiment classification [14]. Traditional sentiment analysis typically focuses on coarse-grained polarity judgments and struggles to handle complex contextual phenomena [15]. Aspect-based sentiment analysis has shifted research toward fine-grained opinion extraction. However, its core tasks remain centered on aspect extraction and polarity classification, making it less suited to constructing continuous variables for design optimization [16]. By contrast, Transformer-based large language models have stronger contextual understanding and instruction-following capabilities, enabling them to process multidimensional sentiment information [16]. Nevertheless, their performance on complex aspect-level sentiment tasks is constrained by task structure and model boundaries; therefore, they cannot directly replace a stable, comparable framework for design attribution [17]. For this study, the key task is to transform dispersed, colloquial, and implicit experience expressions in user reviews into fixed-dimensional design representations, thereby supporting subsequent satisfaction modeling and design interpretation [6,18].

1.2. Research Gaps

Existing studies have made substantial progress in EV charging research, user review mining, and data-driven design. However, these three research streams have not yet been sufficiently integrated regarding home EV chargers as a specific product system [6,9,19]. First, existing studies have mainly examined charging facilities from the perspectives of infrastructure distribution, market expansion, user behavior, and system scheduling [20]. At the same time, relatively limited attention has been paid to residential EV chargers as product systems requiring design optimization. Second, although user review studies have demonstrated their value in reconstructing user experience, extracting product features, and generating design inspiration, most remain focused on topic identification, requirement summarization, or innovation opportunity discovery. The process of transforming raw review topics into design-dimension variables remains insufficiently developed [21]. Third, although large language models and explainable machine learning have been applied to related analytical tasks and the extraction of design guidance [16], an integrated framework combining topic extraction and explainable machine learning remains lacking for residential EV charger products [22]. In addition, existing data-driven design studies have largely emphasized attribute-importance rankings, with limited attention to asymmetric predictive patterns in which low attribute values are associated with pronounced penalties while high values yield comparatively modest gains [23,24].

1.3. Research Objectives

Based on the above research gaps, this study addresses the following three research questions:

RQ1: What core design dimensions can be identified from user reviews of home EV chargers, and how can raw review topics be further transformed into analytical variables for design research?

RQ2: Can stable predictive associations be established between semantic proxy variables for the design dimensions and platform-rated satisfaction, and do these associations differ across design dimensions?

RQ3: How do different design dimensions differ in predictive importance, SHAP attribution direction, and interval-specific association patterns, and how can these patterns inform design optimization priorities for home EV chargers?

This study makes three contributions. First, it conceptualizes residential EV chargers as household product systems and identifies the design dimensions shaping residential charging experience [25]. Second, it develops a review-mining framework integrating BERTopic-based topic identification, validated design-dimension consolidation, evidence-constrained aspect-level coding, and machine learning prediction, thereby transforming unstructured reviews into fixed-dimensional design variables with explicit validation [26,27]. Third, through model comparison, SHAP interpretation, and robustness checks, it identifies design dimensions consistently associated with platform-rated satisfaction and translates these associations into product-level optimization priorities [28]. The findings provide evidence-informed guidance for residential charger development, subject to the limitations of platform ratings and single-platform data.

2. Research Framework and Methods

2.1. Research Framework

This study develops a research framework that integrates review collection and cleaning, BERTopic-based topic identification [22], manual consolidation of design dimensions, large-language-model-based aspect-level continuous coding under structured prompt constraints, machine learning prediction, and SHAP-based interpretive analysis [29]. Specif-

ically, user reviews of home EV chargers were first collected from the JD e-commerce platform and then subjected to preliminary cleaning procedures, including duplicate removal, noise reduction, and filtering for invalid text. BERTopic was then applied to the review corpus to identify raw topics, which were manually consolidated into ten core design dimensions. Within the predefined design-dimension framework, DeepSeek-V3.2 was used under structured prompt constraints to perform continuous aspect-level coding for each review.

The aspect-level scores were represented as semantic proxy variables, while platform star ratings were used as the prediction target for platform-rated satisfaction. For the best-performing model, parameter tuning and early stopping were applied before examining nonlinear predictive associations between the design dimensions and platform-rated satisfaction. Finally, SHAP was used to interpret the fitted model, describing the direction, relative predictive importance, and response patterns of different design dimensions in the satisfaction prediction task. Human validation, an absence-as-zero coding specification, repeated train–test splits, and cross-model SHAP comparisons were used to assess coding reliability, coding-specification sensitivity, partition stability, and ranking stability, respectively. Prospect theory was used only as an interpretive lens for the observed asymmetric predictive patterns [30].

The detailed workflow is shown in Figure 1.

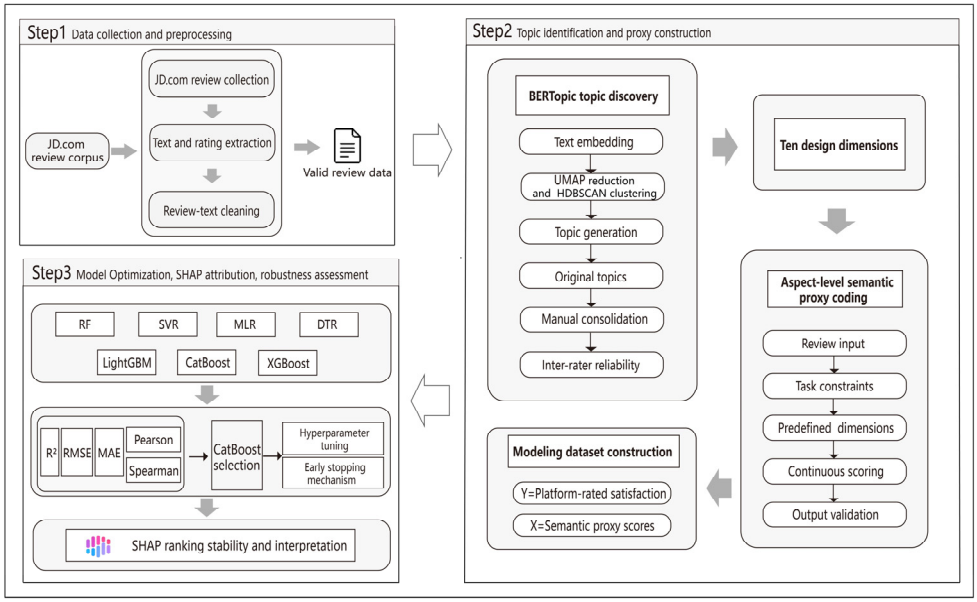


Figure 1. Research framework and analytical workflow.

2.2. Data Sources and Compliance Statement

The review-level dataset was collected in March 2025 from publicly available reviews of home EV chargers on JD.com, one of the major comprehensive e-commerce platforms in China [31,32]. The unit of analysis was the individual review. The raw dataset contained 26,763 records; after removing empty texts, symbol-only entries, punctuation-only entries, exact duplicates, and non-informative entries without identifiable use-related content, 23,893 valid reviews were retained.

This study focused on review-level design-related factors associated with satisfaction outcomes rather than brand- or product-specific comparisons. Because product- or brand-level metadata were not retained, product- or brand-level heterogeneity was not further examined. The data were obtained from publicly accessible platform reviews, did not involve the collection of private information or the identification of individual users, and

were used only for academic research, design-problem diagnosis, pattern summarization, and methodological analysis [33,34].

2.3. BERTopic-Based Topic Identification

2.3.1. Lightweight Text Cleaning

User reviews are typical short-text data characterized by fragmented expression, colloquial language, and substantial noise [35,36]. To reduce the interference from colloquial expressions, fragmented terms, and character-level noise in topic identification, the original review corpus was lightly preprocessed before BERTopic modeling [37]. First, a small number of abnormally long reviews were truncated to 512 characters as an engineering control to reduce the influence of extreme text length on tokenization and embedding consistency [38,39]. Chinese word segmentation was then performed using a domain-specific lexicon. Public Chinese stopword lists were further manually reviewed and supplemented based on the review corpus to remove punctuation, common function words, semantically empty characters, and low-information terms unrelated to topic identification [40,41]. Finally, invalid terms with fewer than 2 characters, purely numerical entries, and blank samples were eliminated. To ensure consistency in subsequent modeling, the row-wise correspondence between the segmented texts and original reviews was retained during the removal of invalid rows [38].

2.3.2. Text Embedding Representation

Traditional topic models such as LDA mainly rely on word-frequency co-occurrence patterns and have limited capacity to capture semantic compression, synonymic variation, and insufficient contextual information in short texts [37]. To improve the semantic representation of user reviews, this study used a BERTopic-based embedding approach to map review texts into a semantic vector space [42]. Specifically, the pretrained paraphrase-multilingual-MiniLM-L12-v2 model from the sentence-transformers framework was used to generate sentence-level vector representations of the review texts. The resulting embeddings were then fed into BERTopic for subsequent dimensionality reduction, clustering, and topic generation [43].

Let $d_i^{'}$ denote the i-th valid review. Its embedding representation is defined as follows:

$$\mathbf{e}_i = f\left(d_i^{'}\right) \tag{1}$$

Here, $f(\cdot)$ denotes the text encoding function, and $\mathbf{e}_i$ represents the semantic embedding vector of the review. Compared with traditional bag-of-words representations, this approach better preserves semantic relationships among reviews, allowing reviews with different expressions but similar concerns to be appropriately grouped in the vector space [44].

2.3.3. Dimensionality Reduction, Clustering, and Topic Generation

After obtaining the embedding vectors, UMAP was used to reduce the dimensionality of the high-dimensional embeddings, and the reduced representations were then fed into HDBSCAN for density-based clustering [42,45]. UMAP is a topology-based nonlinear dimensionality reduction method that effectively preserves local neighborhood structures, making it suitable for processing high-dimensional short-text data [46]. The UMAP parameters were set as follows: n_neighbors = 5, n_components = 6, min_dist = 0.0, metric = “cosine”, and random_state = 42.

After dimensionality reduction, HDBSCAN was used for clustering with the following parameters: min_cluster_size = 60, min_samples = 3, and metric = “Euclidean”. This configuration helped limit the generation of overly small topic clusters while retaining the ability

to identify noisy samples [47]. Based on these steps, BERTopic integrated precomputed embeddings, a vectorization model, UMAP-based dimensionality reduction, and HDBSCAN clustering results to identify topics automatically. The parameter nr_topics = "auto" was used, allowing the model to adjust the number of topics based on the clustering results [48].

After topic generation, class-based TF-IDF, namely c-TF-IDF, was used to represent each topic [48,49]. It is calculated as follows:

$$W_{t,c} = tf_{t,c} \cdot \log\left(1 + \frac{A}{tf_t}\right) \quad (2)$$

where $W_{t,c}$ denotes the weight of term t in topic c, $tf_{t,c}$ denotes the frequency of term t within topic c, tf_t denotes the total frequency of term t across all topics, and A denotes the size of the overall corpus.

2.3.4. Manual Consolidation of BERTopic-Derived Topics

The BERTopic-derived topic clusters were consolidated into design-oriented dimensions for subsequent aspect-level coding rather than being used directly as model inputs. Two researchers independently assigned each valid topic to one design dimension. Intercoder reliability was assessed using percent agreement and Cohen's κ. A third researcher reviewed discrepant assignments and facilitated consensus resolution.

2.4. LLM-Based Aspect-Level Continuous Coding and Human Validation

Based on the ten design dimensions derived from topic identification and manual consolidation, this study used DeepSeek-V3.2 under structured prompt constraints to perform aspect-level continuous coding for each review [50]. Under the predefined design-dimension framework, each review was mapped to a ten-dimensional continuous semantic vector for subsequent structured modeling.

Given the requirements of Chinese short-text understanding, instruction following, and structured output generation, this study used the DeepSeek API to perform continuous aspect-level coding of review texts. The DeepSeek-V3 series has achieved scores of 86.1, 86.5, 90.9, and 64.8 on IF-Eval, C-Eval, CLUEWSC, and C-SimpleQA, respectively, indicating its suitability for instruction-following and Chinese semantic-processing tasks [51]. Compared with simple polarity labels, continuous representations can better preserve dimensional differences in reviews and provide richer information for subsequent regression modeling and nonlinear interpretation [52].

The LLM was guided by a structured prompt that specified the review input, task constraints, design-dimension definitions, continuous scoring rules, and fixed-format output requirements. Although LLMs can support text annotation and semantic coding, their outputs require validation rather than being treated as ground-truth labels. Accordingly, the LLM-generated scores in this study were used as semantic proxy variables, with explicit task definitions, constrained outputs, prompt transparency, output validation, and comparison with human judgments [53]. Figure 2 illustrates the structured prompting framework. Through this protocol, the model was constrained to perform coding within predefined design-semantic boundaries [50].

Let R = {r_1, r_2, …, r_N} denote the review corpus and A = {a1, a2, …, a10} the predefined design dimensions. For each review r_i, the prompting function $P(\cdot)$ embedded the original text into a fixed task template. The LLM coding function $f(\cdot)$ then generated the corresponding aspect-level vector:

$$\mathbf{x}_i = f(P(r_i)) = [s_{i1}, s_{i2}, \ldots, s_{i10}], \quad s_{ij} \in [-1, 1]. \quad (3)$$

Here, s_{ij} denotes the semantic proxy score of review i for design dimension j. Its sign represents evaluative direction, whereas its absolute value represents the intensity of dimension-specific semantic evidence [54]. For each review, the continuous coding values across the ten design dimensions were constrained within the interval [−1, 1].

The generated variables were treated as structured semantic proxy variables rather than direct measurements of users' latent attitudes [55].

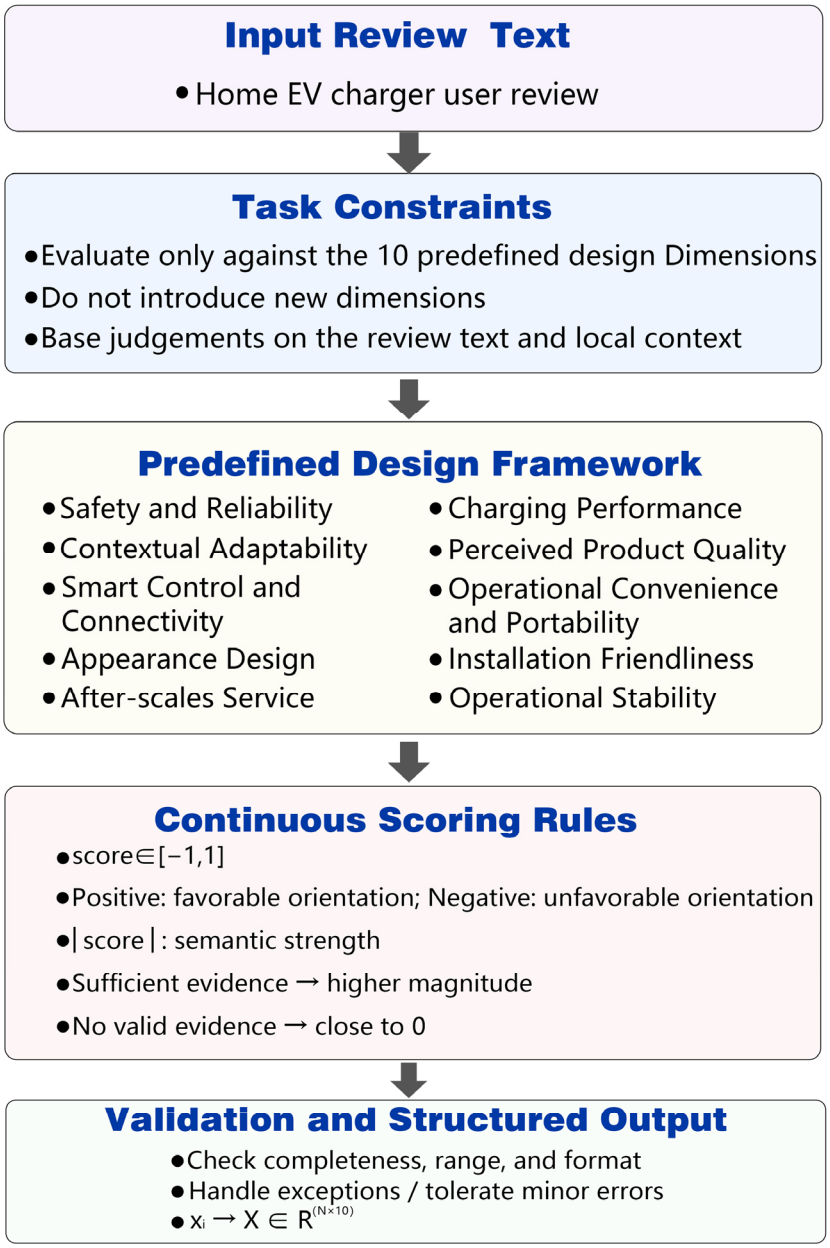


Figure 2. Structured prompting framework for aspect-level continuous coding of design dimensions using deepSeek.

In the primary coding scheme, explicit textual evidence supported medium- or high-magnitude scores, whereas limited contextual cues were restricted to low-magnitude scores. To test whether such values affected model interpretation, a matched absence-as-zero matrix was constructed by assigning 0 to review–dimension units without explicit or contextually inferable dimension-specific evidence. The alternative matrix used the same samples, score range, CatBoost configuration, and test partition as the primary matrix and served exclusively as a coding-specification sensitivity analysis.

After processing, all reviews can be further represented as the following feature matrix:

$$X = \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1,10} \\ s_{21} & s_{22} & \cdots & s_{2,10} \\ \vdots & \vdots & \ddots & \vdots \\ s_{N1} & s_{N2} & \cdots & s_{N,10} \end{bmatrix} \in \mathbb{R}^{N\times 10} \quad (4)$$

Each row of X corresponds to the coding result of a review, and each column corresponds to a fixed design dimension. Together, they constitute the input feature space for the satisfaction prediction model. To ensure the machine readability of the coding outputs, a fixed-field output format was specified, and programmatic validation was performed [56]. If missing fields, out-of-range values, or format-parsing failures were detected, retry procedures or exception-tolerant handling were triggered.

Human Validation of LLM-Generated Aspect-Level Scores

To evaluate the reliability of the LLM-generated aspect-level scores, a stratified subset of 400 reviews was manually annotated, yielding 4000 review–dimension units. Each unit was coded for evidence type, polarity, and semantic intensity. Units without valid evidence or containing only neutral factual statements were assigned 0; contextual cues were restricted to low-intensity scores, whereas explicit evidence could support the full intensity range. Manual scores were mapped to the same $[-1,1]$ scale as the LLM outputs, with polarity assessed at the dimension level rather than inferred from the overall review tone. Inter-coder reliability was evaluated using agreement rates, Cohen's κ, weighted κ, Pearson correlation, and mean absolute error.

2.5. Machine Learning Interpretation Models and Comparison Strategy

After obtaining the ten-dimensional semantic proxy scores for each review, platform star ratings were used as the dependent variable for predicting platform-rated satisfaction [57]. Platform ratings were introduced only after feature construction and were unavailable to both the LLM and human coders, thereby preventing direct target leakage. Because review text and ratings originated from the same evaluative process, semantic overlap could not be excluded; the matched absence-as-zero analysis therefore examined whether the principal SHAP findings depended on weak values assigned to unmentioned dimensions.

Because conventional linear regression models are typically based on strict linearity assumptions and may fail to capture complex nonlinear mappings and higher-order interactions among variables, machine learning regression methods were used to analyze the relationships between design dimensions and user satisfaction [58,59]. Rather than assuming that any specific model would be optimal, this study compared seven models: linear regression, support vector regression, decision tree regression, random forests, XGBoost, LightGBM, and CatBoost. This comparison was conducted to identify the best-performing method for the present research context more objectively [60].

To reduce the randomness introduced by a single train–test split, a repeated random-splitting strategy was adopted for model comparison. The coefficient of determination (R^2) was used as the primary metric for evaluating overall model performance [61,62]. Here, R^2 measures the extent to which the model explains the variance in user ratings and is calculated as follows [63]:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad (5)$$

where y_i denotes the observed rating, $\hat{y}_i$ denotes the predicted rating, and $\bar{y}$ denotes the mean observed rating. R^2 was used to measure the extent to which the model explained

variance in the ratings, whereas MAE, MSE, and RMSE were used to evaluate prediction errors. Pearson and Spearman correlation coefficients were also used to examine the consistency between predicted and observed values [64,65]. The best-performing model was ultimately selected based on the comprehensive comparison results.

Based on the subsequent comparison results, lightweight hyperparameter optimization was further applied to the best-performing model to improve its generalization ability and stability [66]. After model comparison using repeated random splitting, the dataset was further split into fixed training and test sets at an 8:2 ratio. The training set was used for model tuning and fitting, whereas the test set was used for final performance evaluation and subsequent SHAP-based interpretive analysis. Hyperparameter optimization was performed within a moderate search space covering key parameters: iterations = 100, 200, 300, and 500; learning_rate = 0.03, 0.05, 0.08, and 0.10; depth = 4, 6, 8, and 10; and l2_leaf_reg = 1, 3, 5, and 7. RandomizedSearchCV with three-fold cross-validation was used for the sampled search, with R^2 as the primary comparison metric [67,68].

In addition, an early-stopping mechanism was introduced by splitting the training set into a validation subset and using validation performance as the criterion for terminating iterations. This strategy was used to reduce overfitting to local noise during later training stages [69,70]. Finally, SHAP-based interpretive analysis was conducted on the model selected and optimized using the procedures above.

As a supplementary robustness check, the selected feature matrix was further evaluated under repeated data partitions. Specifically, seven regression models were compared across 20 repeated stratified train–test splits with different random seeds. For each model, R^2, MAE, RMSE, Pearson correlation, and Spearman correlation were reported as mean ± standard deviation.

To examine whether the SHAP interpretation depended on the selected model specification, importance rankings were compared across five tree-based ensemble models: CatBoost, XGBoost, LightGBM, random forests, and gradient boosting regression. Multiple linear regression and support vector regression were excluded to maintain consistency in tree-based SHAP estimation and feature-ranking comparability. Ranking stability was evaluated using Top-5 frequency and Spearman rank correlation relative to CatBoost.

2.6. SHAP-Based Interpretive Analysis

To examine the attribution direction, magnitude, and relative predictive importance of different design dimensions in the platform-rated satisfaction prediction model, this study applied SHAP-based attribution analysis to the best-performing machine learning model [71]. SHAP is a feature attribution method based on cooperative game theory. Its core idea is to decompose the prediction for an individual sample into the sum of a baseline output and the contributions of individual features within an additive explanation framework [72], as follows:

$$g\left(x^{'}\right) = \phi_0 + \sum_{m=1}^{M} \phi_m x_m^{'} \tag{6}$$

where $g\left(x^{'}\right)$ denotes the explanation model, ϕ_0 denotes the model output under the baseline condition, M denotes the number of input features, $x_m^{'}$ denotes the m-th simplified feature variable, and ϕ_m denotes the contribution of this feature to the model prediction [72]. The theoretical foundation of SHAP is derived from the Shapley value. This method determines the final attribution value of a feature by evaluating its marginal contribution

across different feature coalitions [73]. For a feature m, its Shapley value can be expressed as follows:

$$\phi_m(f,x) = \sum_{S \subseteq F \setminus \{m\}} \frac{|S|!(|F|-|S|-1)!}{|F|!} \left[f_{S\cup\{m\}}\left(x_{S\cup\{m\}}\right) - f_S(x_S) \right] \quad (7)$$

where F denotes the full set of features, S denotes an arbitrary subset of features that does not include feature m, and $|S|$ denotes the number of features in subset S [74]. $f_S(x_S)$ represents the expected model output when only the feature subset S is given. This method provides the basis for identifying high-importance dimensions and potential nonlinear attribution patterns. In this study, SHAP values were interpreted as model-based predictive attributions under the constructed feature system, not as evidence of causal effects.

3. Results

3.1. Dataset Descriptive Analysis

To characterize the rating structure of the source corpus, Figure 3 reports the platform ratings of all 26,763 records before text cleaning. Subsequent modeling used the 23,893 valid reviews retained after cleaning. Among the 26,763 user reviews of residential EV chargers, five-star reviews accounted for 81.02%, three-star reviews for 14.82%, one-star reviews for 3.63%, two-star reviews for 0.40%, and four-star reviews for 0.13%. This rating pattern is consistent with the self-selection effect commonly observed in online reviews, whereby users with more salient or polarized experiences are more likely to post reviews [75]. Therefore, this rating structure provides important contextual data for interpreting subsequent prediction results and SHAP attribution patterns.

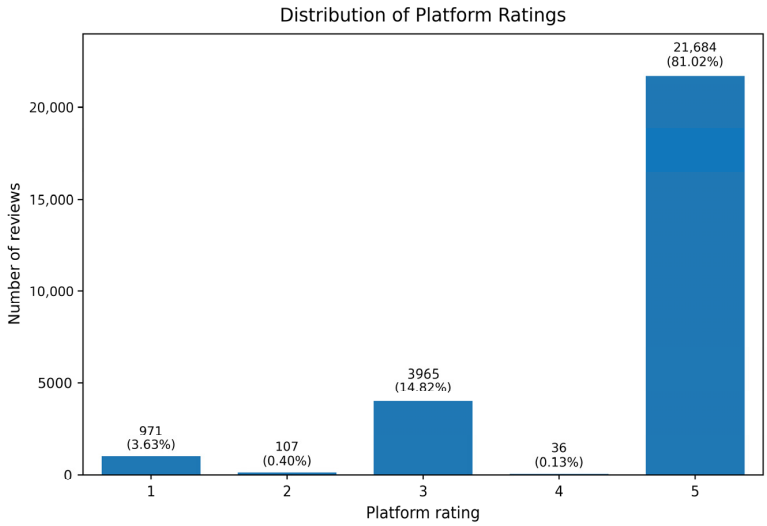


Figure 3. Distribution of platform ratings in the raw collected corpus (n = 26,763).

3.2. BERTopic-Derived Topics and Consolidation Reliability

After lightweight text cleaning, BERTopic was applied to the original review data to identify raw topics. The results produced an initial topic structure that included an outlier class. Excluding the outlier class −1, 16 valid clustered topics were identified, as detailed in Table 1.

The 16 topics were mapped onto 10 design dimensions based on semantic coherence, design-object specificity, dominant semantic orientation, and modeling operability. The two independent coders agreed on 12 topics, yielding a percent agreement of 75.00% and a Cohen's κ of 0.72, indicating substantial inter-coder agreement. The remaining four topics were reviewed by a third researcher and resolved through consensus discussion. Table 1

presents the topic keywords, topic summaries, and corresponding design dimensions, forming the basis for the subsequent ten-dimensional coding framework.

Table 1. Manual consolidation of BERTopic-derived topics into the ten design dimensions.

Topic Keywords	Original Topic Summary	Final Design Dimensions
Installation, Charging, Charging station, Professional	Installation implementation	Installation Friendliness
Performance, Safety, Protection, Level	Charging Performance and safety protection	Charging Performance
Interface, Post-operation, Normal, Use	Usage process and post-operation	Operational Stability
Portability, Operation, Convenient, Handiness	Operation and portability experience	Operational Convenience and Portability
Use, Product, Mark, A period of time	Sustained product use experience	Perceived Product Quality
Awesome, Simple, Great, Huge	Convenience and positive experience	Operational Convenience and Portability
Portability, Carry, Trunk, Convenient	Carrying and portability experience	Operational Convenience and Portability
Rain, Rainproof, Waterproof, Rainy days	Environmental protection adaptability	Contextual Adaptability
Appearance design, Good-looking, Appearance, Beautiful	Aesthetic appearance	Appearance Design
Integrated, Delivery and installation, Worry-free, Xiaomi	Integrated delivery– installation coordination	Installation Friendliness
Password lock, Password, Number, Rotate	Safety control and protection	Safety and Reliability
Installation, End of the month, Effect, Install	Installation result feedback	Installation Friendliness
Signal, Mobile phone, Connection, Bluetooth	Device connection and communication	Smart Control and Connectivity
Remote, Control, Function, Convenient	Remote control and intelligent operation	Smart Control and Connectivity
Appearance design, Minimalist, Attractive, Simple	Minimalist styling and visual perception	Appearance Design
Shopping, Experience, Dissuade, Deceive	Shopping and service feedback	After-sales Service

3.3. Construction and Validation of Aspect-Level Coding Variables

After the lightweight preprocessing for BERTopic modeling described in Section 2.3.1, the original reviews were further cleaned to construct the modeling dataset. Empty texts, symbol-only entries, punctuation-only entries, exact duplicate reviews, and meaningless noise were removed, resulting in 23,893 valid reviews. Based on the predefined 10 design dimensions, LLM-based aspect-level coding was performed to construct a 10-dimensional structured feature matrix. A stratified human validation was conducted on 400 reviews.The construction of the modeling dataset and the human validation sample are summarized in Table 2.

The validation results indicate acceptable evidence-level reliability and positive LLM–human convergence. Evidence identification achieved an agreement rate of 0.8758 and Cohen's $\kappa = 0.6651$, while dimension mention identification reached Cohen's $\kappa = 0.7546$. The LLM-generated scores were positively associated with the human reference scores (Pearson $r = 0.7438$; MAE = 0.3850), and directional consistency reached 0.9300 among non-neutral cases.

Table 2. Human inter-coder reliability and LLM–human score convergence.

Validation Focus	Metric	Value
Evidence identification	Agreement rate/Cohen's κ	0.8758/0.6651
Dimension mention identification	Cohen's κ	0.7546
Polarity and intensity reliability	Weighted κ	0.5942/0.3775
Human reference-score consistency	Pearson r/MAE	0.7712/0.2842
LLM–human score convergence	Pearson r/MAE	0.7438/0.3850
LLM–human directional consistency	Sign accuracy, all/non-neutral cases	0.7440/0.9300

Under the absence-as-zero specification, 73.72% of the 238,930 review–dimension units were classified as absent, 5.55% as implicitly evidenced, and 20.73% as explicitly evidenced. Evidence-supported non-zero scores therefore accounted for 26.28% of all units and ranged from 6.26% for Smart Control and Connectivity to 60.95% for Installation Friendliness. The detailed distribution of evidence types and non-zero scores across the 10 design dimensions is presented in Table 3. Overall, the results indicate acceptable reliability for evidence identification and evaluative direction.

Table 3. Evidence coverage used for the presence-constrained sensitivity analysis.

Design Dimension	Absent %	Implicit %	Explicit %	Non-Zero %
Safety and Reliability	90.32	3.20	6.48	9.68
Charging Performance	74.63	2.98	22.39	25.37
Contextual Adaptability	88.43	7.56	4.01	11.57
Perceived Product Quality	71.23	12.61	16.16	28.77
Smart Control and Connectivity	93.74	1.83	4.43	6.26
Operational Convenience and Portability	56.27	8.45	35.28	43.73
Appearance Design	77.73	2.68	19.60	22.27
Installation Friendliness	39.05	1.77	59.17	60.95
After-sales Service	66.44	1.15	32.41	33.56
Operational Stability	79.35	13.27	7.38	20.65

This evidence-constrained matrix was subsequently compared with the primary coding specification to assess whether low-magnitude scores assigned to weakly supported dimensions materially affected the SHAP interpretation. Figure 4 shows the mean aspect-level coding values across platform rating groups. Low-rating groups generally exhibit negative values, whereas high-rating groups shift toward positive values, indicating directional consistency between the coded variables and platform-rated satisfaction.

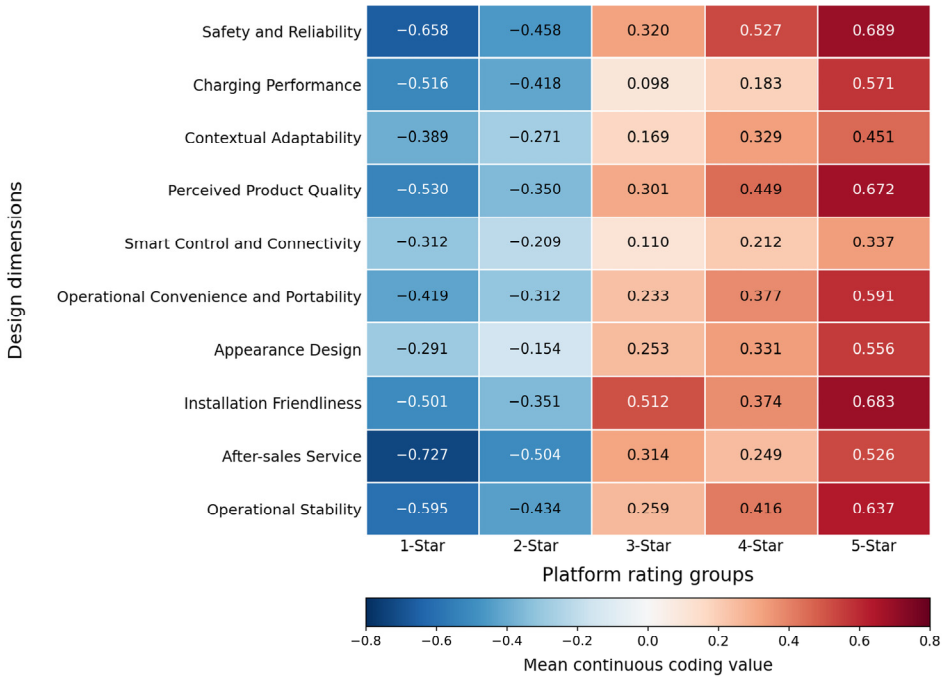


Figure 4. Rating-stratified mean semantic proxy scores across the design dimensions.

3.4. Comparison Results of Machine Learning Models

Table 4 presents the overall performance of the seven machine learning models in the platform-rated satisfaction prediction task. CatBoost achieved the highest R^2 (0.6429), the lowest RMSE (0.6047), and the highest Pearson (0.8019) and Spearman (0.5894) correlations. SVR achieved the lowest MAE (0.3263), indicating that no single model was optimal across all metrics. Compared with XGBoost, LightGBM, random forests, and other models, CatBoost showed a better balance among explanatory capacity, error control, and stability of results. Therefore, CatBoost was selected as the base model for subsequent hyperparameter optimization and SHAP-based interpretive analysis [76]. Because several ensemble models showed similar performance, repeated-split and cross-model analyses were subsequently conducted as robustness checks.

Table 4. Performance comparison of different machine learning models.

Model	R^2	MAE	RMSE	Pearson	Spearman
CatBoost	0.6429	0.3532	0.6047	0.8019	0.5894
XGBoost	0.6383	0.3542	0.6086	0.7991	0.5876
RF	0.6034	0.3653	0.6373	0.7784	0.5776
LightGBM	0.6373	0.3504	0.6095	0.7985	0.5873
SVR	0.6001	0.3263	0.6399	0.7812	0.5456
MLR	0.6087	0.4382	0.6329	0.7804	0.5760
DTR	0.2721	0.3724	0.8633	0.6369	0.5337

3.5. Coding-Specification Sensitivity and SHAP Ranking Stability

As an auxiliary robustness validation procedure, this study evaluated the model's predictive performance via 20 repeated stratified train-test splits. This analysis was intended to assess whether the primary model selection was sensitive to a single data partition. CatBoost showed stable performance across repeated splits, with $R^2 = 0.6425 \pm 0.0078$, RMSE = 0.6022 ± 0.0065, and Pearson correlation = 0.8018 ± 0.0048. These results support the use of CatBoost for the primary SHAP analysis and indicate that approximately 35.75% of the rating variance remained unexplained.The robustness results are summarized in Table 5.

Table 5. Cross-model stability of SHAP importance rankings.

Design Dimension	Mean Rank	Top-5 Frequency
Perceived Product Quality	1.0	5/5
Charging Performance	2.6	5/5
Operational Stability	2.8	5/5
Operational Convenience and Portability	3.6	5/5
After-sales Service	5.0	5/5
Contextual Adaptability	6.4	0/5
Installation Friendliness	6.6	0/5
Safety and Reliability	8.4	0/5
Smart Control and Connectivity	8.8	0/5
Appearance Design	9.8	0/5

Using the same CatBoost configuration and test set (n = 4779), the absence-as-zero specification retained four of the five primary Top-5 dimensions. The highest-ranked dimensions were Operational Convenience and Portability (21.66%), Perceived Product Quality (17.40%), After-sales Service (13.84%), Charging Performance (12.91%), and Appearance Design (11.98%). The overall directional SHAP pattern was preserved, but Operational Stability decreased in rank, whereas Appearance Design increased. The comparison there-

fore supports partial stability of the high-importance set rather than invariance of the exact ranking.

Across the five ensemble models, Perceived Product Quality, Charging Performance, Operational Stability, Operational Convenience and Portability, and After-sales Service appeared in every Top-5. Within the robustness analysis, Spearman rank correlations with the reference CatBoost ranking ranged from 0.9273 to 0.9758 (all $p < 0.001$).

The primary SHAP-based importance ranking was obtained from the final CatBoost model reported in Figure 5. Cross-model comparison was subsequently conducted as a supplementary robustness analysis to assess whether the set of highest-importance dimensions remained consistent across alternative tree-based ensemble models. Charging Performance, Operational Stability, Perceived Product Quality, Operational Convenience and Portability, and After-sales Service consistently ranked among the Top-5 features across all five models, indicating substantial model-level robustness of the principal importance pattern. The consistency observed under the alternative absence-as-zero specification further indicates that low-magnitude continuous coding did not compromise the robustness of the principal findings, while preserving weak or implicit evaluative information that would otherwise be lost under strict zero assignment.

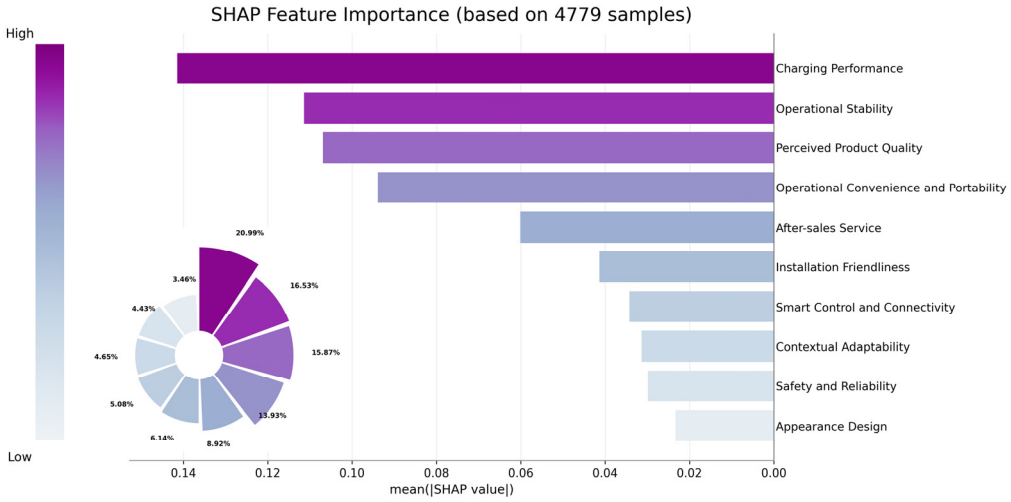


Figure 5. Global SHAP importance ranking of design dimensions.

4. SHAP-Based Predictive Interpretation of Design Dimensions

4.1. Global Importance and Beeswarm Plot Analysis

The SHAP results showed substantial variation in the predictive importance of different design dimensions in the CatBoost model. As shown in Figure 5, Charging Performance had the highest mean absolute SHAP value, accounting for 20.99% of the total importance, indicating that this dimension had the strongest overall contribution to the model prediction. It was followed by Operational Stability (16.53%), Perceived Product Quality (15.87%), and Operational Convenience and Portability (13.93%). These four dimensions were the main coded features distinguishing between higher- and lower-rated predictions in the primary CatBoost model.

By contrast, After-sales Service accounted for 8.92% of the total importance, still showing a relatively stable explanatory role in the model. Installation Friendliness (6.14%), Smart Control and Connectivity (5.08%), Contextual Adaptability (4.65%), Safety and Reliability (4.43%), and Appearance Design (3.46%) showed relatively lower average contributions. Overall, design dimensions directly related to the core use process showed higher predictive importance. In contrast, appearance, smart connectivity, and several extended dimensions made relatively limited average marginal contributions to satisfaction

prediction. Accordingly, the subsequent interpretation focuses on dimensions that recur within the high-importance set and on stable directional patterns, rather than on exact ranks or percentage differences.

The SHAP beeswarm plot further reveals the effect directions and distributional differences for each design dimension at the sample level [72]. As shown in Figure 6, the feature points for Charging Performance, Operational Stability, Perceived Product Quality, and Operational Convenience and Portability are more widely distributed along the horizontal axis. Samples with higher feature values are more frequently located in the positive SHAP region, whereas samples with lower feature values are more often distributed in the negative SHAP region. These patterns indicate that these core dimensions are major sources of variation in predicted satisfaction across samples.

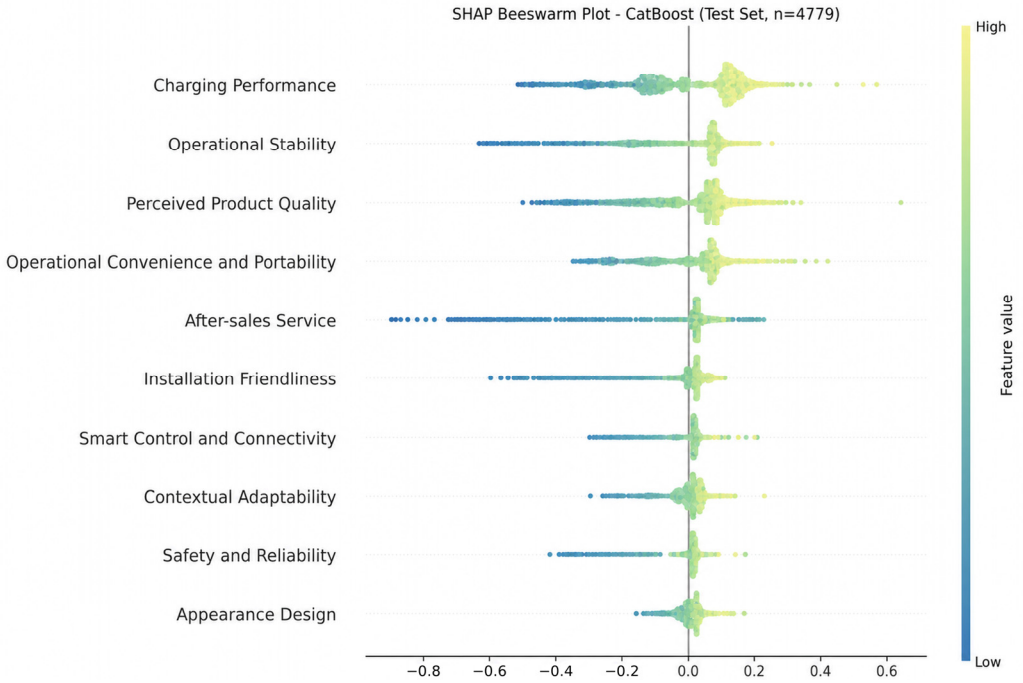


Figure 6. SHAP beeswarm plot of design dimensions.

Among the analyzed dimensions, Charging Performance exhibits the broadest distribution range, while Operational Stability and Perceived Product Quality similarly demonstrate substantial dispersion. In contrast, Operational Convenience and Portability features more distinct directional consistency. Nevertheless, local regions still present color intermixture and overlapping between positive and negative SHAP values, which indicates that the aforementioned relationships do not conform to strict linearity and possess a certain degree of heterogeneity.

By contrast, After-sales Service and Installation Friendliness show more evident low-value penalty patterns. In particular, After-sales Service exhibits the most pronounced tail extension in the negative SHAP region, with some samples approaching −0.9. However, the positive contributions of its high-value samples are mainly concentrated around 0 to 0.2, indicating that this dimension primarily helps prevent a marked decline in predicted satisfaction rather than continuously raising the upper bound of prediction. The remaining point clouds are more concentrated around zero and show narrower overall dispersion. Although Safety and Reliability had relatively low global importance, low feature values still showed a limited negative extension.

4.2. Dependence Plot Analysis

To further examine how model attribution varied across the value ranges of key design dimensions, SHAP dependence plots were generated for the six highest-ranked variables.The resulting dependence plots are presented in Figure 7. These plots were used to describe interval-specific predictive association patterns, including threshold-like transitions, convergence in high-value ranges, and low-value penalties. This section primarily reports the model's revealed dependency patterns.

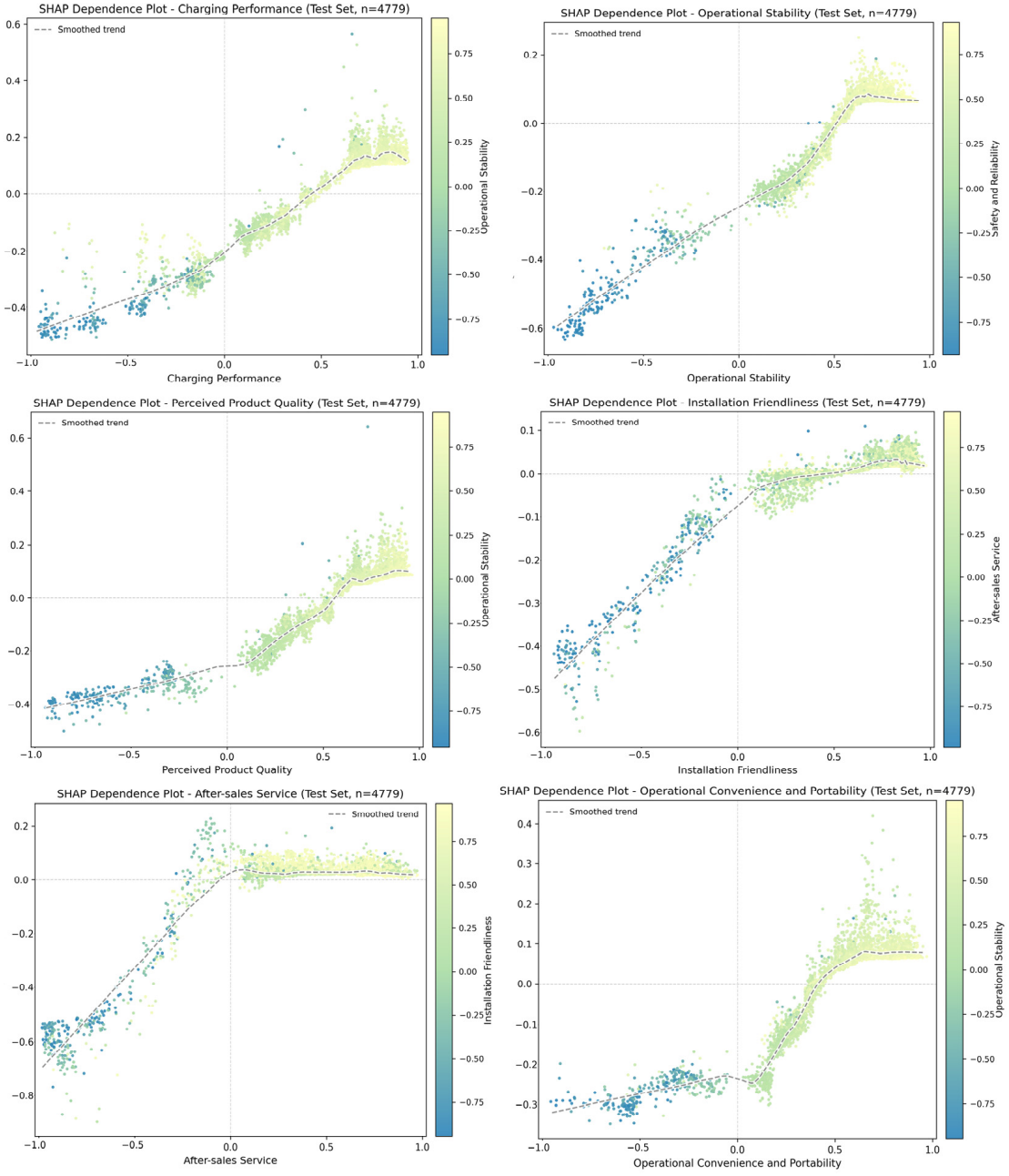


Figure 7. SHAP dependence plots for key design dimensions.

Charging Performance and Operational Stability both exhibit nonlinear marginal response patterns, characterized by a transition from negative to positive SHAP values followed by gradual convergence. Specifically, for Charging Performance, SHAP values increase most markedly within the range of approximately 0.40–0.70 and shift from negative to positive at around 0.45–0.50. When the feature value exceeds 0.70, the slope of the

curve clearly decreases, suggesting that although positive contributions remain, marginal gains begin to diminish. Operational Stability shows a similar pattern. Its SHAP values cross zero at approximately 0.50, increase more rapidly within the range of approximately 0.30–0.60, and gradually converge after the feature value exceeds 0.60. This indicates that the contribution of core performance-related dimensions to satisfaction prediction changes most markedly around their basic threshold ranges. These findings suggest that Charging Performance and Operational Stability exhibit clear basic-threshold characteristics: when they fall below an acceptable level, they are more likely to generate negative SHAP contributions; once they cross the basic threshold, their positive contributions begin to emerge and then gradually converge.

By contrast, Perceived Product Quality, Operational Convenience and Portability more clearly exhibit threshold activation and stage-like increases. Perceived Product Quality is mainly associated with negative contributions in the low and medium-low value ranges. However, its SHAP values continue to rise from the medium-low range, complete a key transition in the 0.50–0.65 range, and become positive at approximately 0.55. Operational Convenience and Portability change substantially within the range of 0.15–0.45 and cross zero at approximately 0.38–0.42. This suggests that these two design dimensions, which are directly related to perceived product maturity and operational experience, generally need to reach a certain threshold before producing more stable positive contributions.

After-sales Service and Installation Friendliness are closer to support-oriented constraint variables. After-sales Service shows relatively strong negative contributions when its value is below -0.20, shifts from negative to positive around zero, and then its marginal contribution tends to level off. Installation Friendliness shows a relatively gradual increasing trend, approaching or crossing zero at approximately 0.40–0.45, with a relatively narrow overall range of variation. This indicates that the main role of support-oriented dimensions is not to continuously raise the upper bound of predicted satisfaction, but to prevent low-level performance from substantially dragging down overall satisfaction evaluations.

Taken together, the dependence plots suggest an asymmetric predictive structure in which threshold-like transitions, diminishing marginal gains, and low-value penalties coexist. Core use-related dimensions tended to generate more stable positive SHAP values after crossing certain value ranges. In contrast, support-oriented dimensions, especially installation and after-sales service, were more likely to produce negative predictive drag when coded at low levels. These patterns provide a model-based basis for the subsequent theoretical discussion.

5. Discussion

5.1. Theoretical Interpretation of Asymmetric Predictive Patterns

The SHAP results reported in Sections 4.1 and 4.2 did not show a simple linear accumulation across all design dimensions. Instead, the attribution structure was characterized by the dominance of core use-related dimensions, low-value penalties in support-oriented dimensions, and diminishing incremental attributions at higher feature values [77]. This study further adopts prospect theory as an explanatory lens to interpret these asymmetric attribution patterns. Prospect theory is used here as an interpretive lens rather than as a directly tested psychological model. According to this theory, evaluations depend on reference points, and experiences below the reference point may be perceived as losses with greater psychological weight than equivalent gains [78].

When applied to the present findings, this theory suggests that stable start-up, continuous charging, controllable installation, recoverable faults, and accessible after-sales support can be considered basic reference conditions for evaluating residential EV chargers. When these conditions are not satisfied, reviews and ratings are more likely to show nega-

tive tendencies. Once core expectations are met, further improvements may still generate positive contributions, but the additional gains tend to diminish. Charging Performance, Operational Stability, Perceived Product Quality, and Operational Convenience and Portability constitute the main explanatory sources of differences in predicted user satisfaction. This indicates that evaluations of residential EV chargers are primarily organized around the use scenario of completing charging in a stable, smooth, and low-effort manner. Among these dimensions, Charging Performance and Operational Stability are closer to basic functional dimensions. In contrast, Perceived Product Quality, Operational Convenience and Portability are more closely related to experience-enhancing dimensions.

It should be noted that although Safety and Reliability does not rank highly in global importance, this does not imply a low design priority for safety [72]. For products such as residential EV chargers, safety is more likely to function as a basic compliance requirement and a bottom-line constraint. Its relatively low mean SHAP contribution may be related to insufficient explicit expression of safety problems in the sample or limited variation in safety-related features [79].

These findings are generally consistent with previous studies that used user reviews and pretrained language models to identify key touchpoints in charging experiences. Such studies similarly suggest that user evaluations are first organized around whether charging is smooth, whether the equipment is reliable, and whether the use process is low-friction [80–82]. The difference is that previous studies have mainly focused on service diagnosis and understanding station operations in public charging contexts. In contrast, this study applies review mining to the design interpretation of residential charging terminals, focusing on how different design dimensions contribute to satisfaction in home charging contexts [81,83,84].

On the other hand, the results do not support treating all design attributes as homogeneous sources of added value. Methodologically, this conclusion echoes the study by Ullah et al., who used SHAP to explain charging-station choice behavior [85]. However, the present study further extends explainable machine learning from the interpretation of public charging choice behavior to the explanation of mechanisms underlying residential design experience. It also complements the attribute-classification approach based on the Kano model and the AHP adopted by Liu et al. [8]. Whereas that approach addresses the ranking and classification of attribute importance, the present study characterizes interval-specific attribution changes, asymmetric patterns, and nonlinear model responses of these attributes across different value intervals and interprets them through prospect theory.

Notably, studies on residential and neighborhood contexts have shown that physical space, electrical capacity, property management, and resident acceptance in multi-unit housing are themselves important barriers to home charging deployment [86]. From this perspective, the more pronounced negative attribution patterns associated with installation and service in this study are not incidental but rather reflect the renewed importance of these design dimensions in residential contexts.

In addition, this study complements existing charging research in terms of how the research object is conceptualized. Existing reviews and systematic studies have long emphasized system-level issues such as infrastructure planning, scheduling control, and dynamic pricing. This line of research is valuable for improving the operational efficiency of public charging networks [87,88]. However, the present findings indicate that, in home-use contexts, users continuously perceive whether product selection, installation, use, fault handling, and after-sales recovery form a recoverable and predictable experience chain. This point is consistent with recent research on residential charging and household energy systems. As home charging is integrated into community-level energy management, dynamic electricity pricing, and smart charging schedules, the importance of design optimization

does not decrease [89]; rather, it becomes more pronounced. This further requires charging terminals to adopt a clearly prioritized and asymmetric allocation of design resources [90].

5.2. Design Optimization Strategies Derived from SHAP-Based Predictive Patterns

Based on the asymmetric predictive patterns identified in the preceding analysis, the design optimization of residential electric vehicle chargers should follow a hierarchical intervention logic. First, manufacturers should prioritize the reliability of the core charging system because low scores for charging initiation, connection continuity, and output stability were associated with pronounced negative SHAP attributions. Second, it is necessary to improve the consistency of perceived product quality and the convenience of low-cost user operation in high-frequency interaction links, as minor user experience defects will gradually accumulate over the long term in residential use scenarios. Third, installation adaptability and after-sales fault recovery capabilities should be integrated into the pre-design phase of the product system, rather than being regarded as complementary compensation for post-purchase services. The following strategies transform the empirical findings derived from the proposed model into product-level design objectives.

5.2.1. Preventing Failure in the Core Charging Chain: Improving Stability and Recovery Support

For optimization of the core use chain, the design focus should shift from merely improving rated parameters to ensuring stable start-up, continuous connection, and recoverable faults. Start-up failures, connection interruptions, and unstable output can weaken users' perceptions of device reliability and process controllability. Therefore, vehicle-recognition logic, connection-confirmation feedback, and start-up response procedures should be optimized so that users can quickly determine whether the device has entered an effective charging state after plugging in the connector.

At the structural level, optimization should focus on the connector-fitting structure, cable-connection structure, interface contact components, and key durability components to reduce looseness, poor contact, communication interruptions, and output fluctuations. For abnormal interruptions, over-temperature protection, or communication failures, clear feedback should be provided through the status display interface, fault-prompt logic, and recovery guidance path, enabling users to identify the nature of the abnormality and take actionable steps. In product development practice, these priorities can be translated into prototype validation criteria such as start-up success rate, reconnection after interruption, connector engagement stability, output fluctuation control, and the clarity of fault-recovery prompts.

5.2.2. Reducing Perceived Loss in High-Frequency Operations: Optimizing Quality Consistency and Low-Effort Interaction

The strategic focus of high-frequency interaction optimization is to reduce the friction that accumulates over time with repeated use. Once core functions reach an acceptable level, experience optimization should shift toward reducing small but cumulative burdens in daily use. The tactile qualities of housing materials, surface treatments, assembly seams, and the consistency of structural integration are key entry points for enhancing perceived product quality. These improvements can reduce detail-level problems that weaken product credibility, such as loose housing, rough interfaces, stiff cables, and unstable brackets.

Meanwhile, status-light signals, icon recognition, prompt wording, and interface information hierarchy should be optimized so that users can accurately identify key states, such as connection, charging, abnormality, and completion, even at night, under low illumination, or from a distance. The grip angle of the charging connector, cable flexibility, bracket position, storage path, and scheduled-charging procedure should also be opti-

mized to reduce burdens such as bending, dragging, cable winding, repeated confirmation, and multi-step navigation, thereby strengthening users' sustained perception of product maturity and usability [91]. For manufacturers, these findings can be operationalized as usability-testing criteria, including one-handed connector operation, night-time state recognition, reduced charging-task steps, convenient cable storage, and user recognition of charging, completion, and abnormal states.

5.2.3. Front-Loading the Control of Installation and After-Sales Risks: Enhancing Adaptability and Recoverability

For installation and support optimization, these elements should be moved from post-purchase service remediation to an integral part of product system design. Installation and after-sales support are usually not value-added attributes that users actively pursue, but bottom-line conditions for whether the product can be successfully deployed and continuously used. When installation conditions are unclear, construction boundaries are ambiguous, or no fault-recovery path is available, users are more likely to perceive these situations as loss events that fall short of their basic expectations [92].

Prior to installation, power supply conditions, spatial requirements, wiring approaches and installation boundaries shall be explicitly confirmed, so as to mitigate issues including infeasible installation after product delivery, mismatched construction conditions and cost overruns. In terms of structural design, appropriate mounting holes, disassembly boundaries, maintenance space and module accessibility shall be reserved to enhance deployment compatibility and maintenance efficiency.

In after-sales and fault-recovery processes, abnormal-state prompts, fault-identification explanations, recovery guidance, and after-sales contact entries should be strengthened to help users quickly identify the source of the problem and obtain a clear resolution path [93]. For problems requiring manual intervention, clear response time limits and progress feedback should be provided to reduce perceived losses during waiting, communication, and the confirmation of responsibility. Only when installation, use, and maintenance form a continuous, traceable, and recoverable support chain can users' judgments of product controllability and service credibility be supported, and the positive value of core performance and interaction experience be reliably realized. The corresponding design strategies are summarized in Table 6.In practice, manufacturers can incorporate these priorities into pre-installation checklists, standardized mounting kits, fault-code explanations, response-time commitments, and visible after-sales progress feedback.

Table 6. Product-level design optimization strategies derived from SHAP-based predictive association patterns.

Strategic Track	Related Dimensions and Predictive Pattern	Optimization Direction	Key Design Targets and Development Examples
Core charging-chain stability optimization	Charging Performance and Operational Stability; threshold-like transition and low-value penalty	Ensure stable start-up and connection continuity	Vehicle recognition; connection confirmation; start-up response; connector engagement; cable connection; output-stability control; prototype tests for start-up success and reconnection after interruption
Fault recovery and risk-response optimization	Charging Performance and Operational Stability, with Safety and Reliability as a baseline constraint; low-value penalty	Strengthen fault recovery support	Status display; fault-prompt logic; overheating-protection feedback; communication-failure recovery; recovery guidance pathway; fault-code differentiation and abnormal-state feedback validation

Table 6. *Cont.*

Strategic Track	Related Dimensions and Predictive Pattern	Optimization Direction	Key Design Targets and Development Examples
Perceived-quality consistency optimization	Perceived Product Quality; threshold-like response and diminishing marginal gain	Enhance consistency of perceived quality	Housing tactility; surface finishing; assembly joints; structural integration consistency; material, surface, assembly gap, and cable durability specifications
Low-burden interaction optimization	Operational Convenience and Portability, supported by Smart Control and Connectivity; reduction in interactional friction	Improve information feedback clarity and support low-burden operation	Status lights; icon recognition; interface hierarchy; charging-connector grip angle; cable flexibility; storage path; scheduled-charging workflow; usability tests for one-hand operation and state recognition
Installation adaptation and service-recovery optimization	Installation Friendliness and After-sales Service; baseline constraint and low-value penalty	Front-load installation adaptation and service recovery	Power-supply and spatial conditions; cable routing; installation boundaries; maintenance accessibility; fault diagnosis; response timeliness; pre-installation checklist, standardized mounting kit, and visible after-sales progress feedback

5.3. Limitations and Future Research Directions

This study has four main limitations. First, user reviews were collected from a single e-commerce platform, and online reviews may be affected by self-selection bias because users with more polarized experiences are more likely to post reviews. Therefore, the results should be interpreted as platform-specific evidence rather than as fully generalizable conclusions for all home EV charger users [94]. Second, product- and brand-level metadata were not retained during data collection; therefore, heterogeneity across brands and product models was not examined. Regional and installation-scenario variables were not collected, limiting the assessment of context-specific variation [95]. The residual variance indicates that the review-derived design variables did not fully capture platform-rated satisfaction. Part of the unexplained variation may stem from product-, transaction-, and installation-related factors not included in the current model, such as price, brand reputation, promotion, logistics, installation cost, product version, and local installation conditions. Platform-specific rating behavior may also contribute to residual prediction errors. Third, the aspect-level codes remain semantic proxies derived from review texts rather than direct measures of users' latent attitudes [96]. Although human validation supported their use in downstream modeling, lower consistency in semantic-intensity coding indicates residual uncertainty for implicit or weakly evidenced expressions. Fourth, the interpretation of the asymmetric predictive patterns was informed by prospect theory and the observed SHAP attribution structure, without direct measurement of users' psychological reference points, loss aversion, or risk perception. Therefore, the related interpretation should be understood as an explanatory discussion based on patterned results rather than as a direct empirical validation of prospect theory.

Future research can proceed in three directions. First, multi-platform and multi-source review data can be introduced to examine whether the identified design priorities remain stable across different markets and platform environments [97]. Second, product-level, brand-level, and installation-scenario variables can be incorporated to improve the control of contextual heterogeneity. Third, interviews and other qualitative methods can be combined to examine how installation, maintenance, and use problems are mapped onto design deficiencies, and to investigate changes in users' risk perception, loss perception, and sense of control during the use of residential EV chargers.

6. Conclusions

To address the difficulty of quantifying user feedback and improving the design of home EV charging devices, this study developed a data-driven design optimization framework that integrates review mining and explainable machine learning. The study generated 10 core design dimensions using BERTopic and manual consolidation [48], introduced a large language model to generate continuous codes, and used platform satisfaction ratings as proxy outcome variables. To strengthen the methodological reliability of the framework, supplementary human validation, absence-as-zero evidence auditing, repeated-split model evaluation, and cross-model SHAP ranking comparisons were added.

The results show that Charging Performance, Operational Stability, Perceived Product Quality, and Operational Convenience and Portability are the core dimensions explaining differences in satisfaction. These findings identify model-based design priorities rather than causal effects. The predictive associations of different design dimensions exhibit clear interval-specific differences and nonlinear characteristics. Interpreted through the lens of prospect theory, the results suggest that design optimization for residential EV chargers should prioritize avoiding touchpoints that users are likely to perceive as critical losses, such as charging failure, uncontrollable installation, unrecoverable faults, and inaccessible after-sales support. This study transforms large-scale review feedback into structured evidence that can be used in design research, providing a context-sensitive quantitative basis for the iteration and functional improvement of residential EV charging devices [98]. It also identifies optimization pathways to improve the home-use EV experience and related service systems at the system level.

Author Contributions: Conceptualization, Y.Z. and K.C.; methodology, Y.Z. and K.C.; software, Y.Z.; validation, Y.Z., Y.P. and J.W.; formal analysis, Y.Z.; investigation, Y.Z.; resources, S.C. and K.C.; data curation, Y.Z.; writing—original draft preparation, Y.Z.; writing—review and editing, Y.P., J.W., Z.L., Y.W., S.C. and K.C.; visualization, Y.Z.; supervision, S.C. and K.C.; project administration, S.C. and K.C.; funding acquisition, S.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the General Project for Humanities and Social Sciences Research, Ministry of Education of China, grant number 23WJA760016.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available from the corresponding author upon reasonable request.

Acknowledgments: During the preparation of this study, the authors used DeepSeek-V3.2 for aspect-level continuous coding of user reviews under a predefined structured prompting framework. The model was used to transform cleaned, unstructured review texts into continuous semantic proxy variables corresponding to ten predefined design dimensions. The outputs were programmatically checked for completeness, format validity, and value-range consistency. The authors reviewed and validated the coding procedure and the interpretation of the results, and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. International Energy Agency. *Global EV Outlook 2025*; International Energy Agency: Paris, France, 2025.
2. White, L.V.; Carrel, A.L.; Shi, W.; Sintov, N.D. Why are charging stations associated with electric vehicle adoption? Untangling effects in three United States metropolitan areas. *Energy Res. Soc. Sci.* **2022**, *89*, 102663. [CrossRef]

3. LaMonaca, S.; Ryan, L. The state of play in electric vehicle charging services: A review of infrastructure provision, players, and policies. *Renew. Sustain. Energy Rev.* **2022**, *154*, 111733. [CrossRef]
4. Fabianek, P.; Madlener, R. Multi-criteria assessment of the user experience at e-vehicle charging stations in Germany. *Transp. Res. Part D Transp. Environ.* **2023**, *121*, 103782. [CrossRef]
5. Ziras, C.; Thingvad, M.; Fog, T.; Yousefi, G.; Weckesser, T. An empirical analysis of electric vehicle charging behavior based on real Danish residential charging data. *Electr. Power Syst. Res.* **2024**, *234*, 110556. [CrossRef]
6. Briard, T.; Jean, C.; Aoussat, A.; Véron, P. Challenges for data-driven design in early physical product design: A scientific and industrial perspective. *Comput. Ind.* **2023**, *145*, 103814. [CrossRef]
7. Hardman, S.; Tal, G. Understanding discontinuance among California's electric vehicle owners. *Nat. Energy* **2021**, *6*, 538–545. [CrossRef]
8. Liu, X. Research on Function Optimization of Electric Vehicle Charging Stations Based on User Demand Analysis: An Empirical Study Using the Kano Model and AHP Method. *Sustainability* **2024**, *16*, 10783. [CrossRef]
9. Yang, B.; Liu, Y.; Liang, Y.; Tang, M. Exploiting user experience from online customer reviews for product design. *Int. J. Inf. Manag.* **2019**, *46*, 173–186. [CrossRef]
10. Qin, J.; Zheng, P.; Wang, X. Product redesign and innovation based on online reviews: A multistage combined search method. *Inf. J. Comput.* **2024**, *36*, 742–765. [CrossRef]
11. An, Y.; Oh, H.; Lee, J. Marketing insights from reviews using topic modeling with BERTopic and deep clustering network. *Appl. Sci.* **2023**, *13*, 9443. [CrossRef]
12. Liu, Y.; Zhang, D.; Wan, Q.; Lin, Z. Innovative mining of user requirements through combined topic modeling and sentiment analysis: An automotive case study. *Sensors* **2025**, *25*, 1731. [CrossRef] [PubMed]
13. Chen, S. Age-appropriate design of smart senior care product APP interface based on deep learning. *Heliyon* **2024**, *10*, e28567. [CrossRef] [PubMed]
14. Mao, Y.; Liu, Q.; Zhang, Y. Sentiment analysis methods, applications, and challenges: A systematic literature review. *J. King Saud Univ. Comput. Inf. Sci.* **2024**, *36*, 102048. [CrossRef]
15. Wankhade, M.; Kulkarni, C.; Rao, A.C.S. A survey on aspect-based sentiment analysis: Tasks, methods, and challenges. *Appl. Soft Comput.* **2024**, *167*, 112249. [CrossRef]
16. Perikos, I.; Diamantopoulos, A. Explainable aspect-based sentiment analysis using transformer models. *Big Data Cogn. Comput.* **2024**, *8*, 141. [CrossRef]
17. Hua, Y.C.; Denny, P.; Wicker, J.; Taskova, K. A systematic review of aspect-based sentiment analysis: Domains, methods, and trends. *Artif. Intell. Rev.* **2024**, *57*, 296. [CrossRef]
18. Joung, J.; Kim, H.M. Explainable neural network-based approach to Kano categorisation of product features from online reviews. *Int. J. Prod. Res.* **2022**, *60*, 7053–7073. [CrossRef]
19. Andrenacci, N.; Valentini, M.P. A Literature Review on the Charging Behaviour of Private Electric Vehicles. *Appl. Sci.* **2023**, *13*, 12877. [CrossRef]
20. Patil, P.; Kazemzadeh, K.; Bansal, P. Integration of charging behavior into infrastructure planning and management of electric vehicles: A systematic review and framework. *Sustain. Cities Soc.* **2023**, *88*, 104265. [CrossRef]
21. Cheng, F.-M.; Wang, J.; Chen, C.; Hu, G.-R.; Cao, Z.-J. Product design improvement method driven by online product reviews. *Sci. Rep.* **2025**, *15*, 10252. [CrossRef] [PubMed]
22. Zhang, J.; Du, Z.; Sun, S.; Wang, S. Perception of customer satisfaction and complaints based on BERTopic and interpretable machine learning: Evidence from hotels in Xi'an. *Curr. Issues Tour.* **2025**, *28*, 3168–3190. [CrossRef]
23. Joung, J.; Kim, H.M. Interpretable Machine Learning-Based Approach for Customer Segmentation for New Product Development from Online Product Reviews. *Int. J. Inf. Manag.* **2023**, *70*, 102641. [CrossRef]
24. Park, S.; Kim, H.M. Extracting Product Design Guidance from Online Reviews: An Explainable Neural Network-Based Approach. *Expert Syst. Appl.* **2024**, *236*, 121357. [CrossRef]
25. Visaria, A.A.; Jensen, A.F.; Thorhauge, M.; Mabit, S.E. User preferences for EV charging, pricing schemes, and charging infrastructure. *Transp. Res. Part A Policy Pract.* **2022**, *165*, 120–143. [CrossRef]
26. Timoshenko, A.; Hauser, J.R. Identifying customer needs from user-generated content. *Mark. Sci.* **2019**, *38*, 1–20. [CrossRef]
27. Zhang, M.; Sun, L.; Li, Y.; Wang, G.A.; He, Z. Using supplementary reviews to improve customer requirement identification and product design development. *J. Manag. Sci. Eng.* **2023**, *8*, 584–597. [CrossRef]
28. Lundberg, S.M.; Erion, G.; Chen, H.; DeGrave, A.; Prutkin, J.M.; Nair, B.; Katz, R.; Himmelfarb, J.; Bansal, N.; Lee, S.-I. From local explanations to global understanding with explainable AI for trees. *Nat. Mach. Intell.* **2020**, *2*, 56–67. [CrossRef] [PubMed]
29. Meng, F.; Jia, Y. Data-Driven Prioritization of User Requirements in Health E-Commerce: An Explainable Machine Learning Study. *J. Theor. Appl. Electron. Commer. Res.* **2026**, *21*, 104. [CrossRef]
30. Cai, L.; Cui, Q.; Yang, Z.; Liu, J.; Zhang, K. Revisiting electric vehicle charging station accessibility: A home and workplace dual-scenario perspective. *J. Transp. Geogr.* **2026**, *130*, 104488. [CrossRef]

31. Du, Q.; Kong, D.; Li, Y.; Ye, K. Customer Ratings and Firm Value: Evidence from Big Data Analysis of Online Consumption in China. *Financ. Res. Lett.* **2025**, *75*, 106867. [CrossRef]
32. Shan, S.; Yang, Y.; Li, C. Which Receives More Attention, Online Review Sentiment or Online Review Rating? Spillover Effect Analysis from JD.com. *Behav. Sci.* **2024**, *14*, 823. [CrossRef] [PubMed]
33. Gliniecka, M. The Ethics of Publicly Available Data Research: A Situated Ethics Framework for Reddit. *Soc. Media Soc.* **2023**, *9*, 20563051231192021. [CrossRef]
34. Zhang, Y.; Fu, J.; Lai, J.; Deng, S.; Guo, Z.; Zhong, C.; Tang, J.; Cao, W.; Wu, Y. Reporting of Ethical Considerations in Qualitative Research Utilizing Social Media Data on Public Health Care: Scoping Review. *J. Med. Internet Res.* **2024**, *26*, e51496. [CrossRef] [PubMed]
35. Fan, Y.; Shi, L.; Yuan, L. Topic modeling methods for short texts: A survey. *J. Intell. Fuzzy Syst.* **2023**, *45*, 1971–1990. [CrossRef]
36. Huo, H.; Yao, J.; Li, Q.; Wu, R.; Wang, C.; Yu, D. Economic vs. ecological benefit for organic food: The role of perceived values in online consumer reviews. *Front. Sustain. Food Syst.* **2025**, *9*, 1469869. [CrossRef]
37. Qiang, J.; Qian, Z.; Li, Y.; Yuan, Y.; Wu, X. Short Text Topic Modeling Techniques, Applications, and Performance: A Survey. *IEEE Trans. Knowl. Data Eng.* **2022**, *34*, 1427–1445. [CrossRef]
38. Siino, M.; Tinnirello, I.; La Cascia, M. Is Text Preprocessing Still Worth the Time? A Comparative Survey on the Influence of Popular Preprocessing Methods on Transformers and Traditional Classifiers. *Inf. Syst.* **2024**, *121*, 102342. [CrossRef]
39. Shen, S.; Zhao, Z.; Yan, D.; Jiang, C.; Wang, D. HsscBERT: Pre-Training Domain Model for the Full Text of Chinese Humanity and Social Science. *npj Herit. Sci.* **2025**, *13*, 229. [CrossRef]
40. Zhang, D.; Li, J.; Xie, Y.; Wulamu, A. Research on Performance Variations of Classifiers with the Influence of Pre-Processing Methods for Chinese Short Text Classification. *PLoS ONE* **2023**, *18*, e0292582. [CrossRef] [PubMed]
41. Chen, W.; Li, Y.; Zhang, X.; Huang, J. Unlocking Visitor Experiences in Cultural Heritage Sites with Chinese Social Media Reviews: A Text Mining Approach. *Herit. Sci.* **2025**, *13*, 414. [CrossRef]
42. Voskergian, D.; Jayousi, R.; Yousef, M. Topic selection for text classification using ensemble topic modeling with grouping, scoring, and modeling approach. *Sci. Rep.* **2024**, *14*, 23516. [CrossRef] [PubMed]
43. Sentence Transformers. Pretrained Models. Available online: https://www.sbert.net/docs/sentence_transformer/pretrained_models.html (accessed on 27 July 2026).
44. Reimers, N.; Gurevych, I. Sentence-BERT: Sentence Embeddings Using Siamese BERT-Networks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing*; Association for Computational Linguistics: Hong Kong, China, 2019; pp. 3982–3992.
45. Feng, Y.; Chen, Z.; Zhang, Y.; Huang, W.; Zhang, X.; He, S. BERTopic_Teen: A Multi-Module Optimization Approach for Short Text Topic Modeling in Adolescent Health. *Front. Public Health* **2025**, *13*, 1608241. [CrossRef] [PubMed]
46. Wang, W.; Wei, F.; Dong, L.; Bao, H.; Yang, N.; Zhou, M. MiniLM: Deep Self-Attention Distillation for Task-Agnostic Compression of Pre-Trained Transformers. In *Advances in Neural Information Processing Systems*; Curran Associates, Inc.: Red Hook, NY, USA, 2020.
47. Campello, R.J.G.B.; Moulavi, D.; Zimek, A.; Sander, J. Hierarchical Density Estimates for Data Clustering, Visualization, and Outlier Detection. *ACM Trans. Knowl. Discov. Data* **2015**, *10*, 5. [CrossRef]
48. Grootendorst, M. BERTopic: Neural Topic Modeling with a Class-Based TF-IDF Procedure. *arXiv* **2022**, arXiv:2203.05794.
49. Li, Y.; Huang, T.; Li, X. Training Path of Big Data Management and Application Talents Based on BERTopic-TOPSIS Model. *PLoS ONE* **2025**, *20*, e0334127. [CrossRef] [PubMed]
50. Lou, R.; Zhang, K.; Yin, W. Large Language Model Instruction Following: A Survey of Progresses and Challenges. *Comput. Linguist.* **2024**, *50*, 1053–1095. [CrossRef]
51. DeepSeek-AI. DeepSeek-V3 Technical Report. *arXiv* **2024**, arXiv:2412.19437.
52. Lee, L.-H.; Yu, L.-C.; Wang, S.; Liao, J. Overview of the SIGHAN 2024 Shared Task for Chinese Dimensional Aspect-Based Sentiment Analysis. In *Proceedings of the 10th SIGHAN Workshop on Chinese Language Processing*; Association for Computational Linguistics: Bangkok, Thailand, 2024; pp. 165–174.
53. Törnberg, P. Best practices for text annotation with large language models. *Sociologica* **2024**, *18*, 67–85. [CrossRef] [PubMed]
54. Taboada, M.; Brooke, J.; Tofiloski, M.; Voll, K.; Stede, M. Lexicon-Based Methods for Sentiment Analysis. *Comput. Linguist.* **2011**, *37*, 267–307. [CrossRef]
55. Humphreys, A.; Wang, R.J.-H. Automated Text Analysis for Consumer Research. *J. Consum. Res.* **2018**, *44*, 1274–1306. [CrossRef]
56. Dagdelen, J.; Dunn, A.; Lee, S.; Walker, N.; Rosen, A.S.; Ceder, G.; Persson, K.A.; Jain, A. Structured Information Extraction from Scientific Text with Large Language Models. *Nat. Commun.* **2024**, *15*, 1418. [CrossRef] [PubMed]
57. Bewick, V.; Cheek, L.; Ball, J. Statistics review 7: Correlation and regression. *Crit. Care* **2003**, *7*, 451–459. [CrossRef] [PubMed]
58. Borisov, V.; Leemann, T.; Seßler, K.; Haug, J.; Pawelczyk, M.; Kasneci, G. Deep Neural Networks and Tabular Data: A Survey. *IEEE Trans. Neural Netw. Learn. Syst.* **2024**, *35*, 7499–7519. [CrossRef] [PubMed]

59. Liang, D.; Frederick, D.A.; Lledo, E.E.; Rosenfield, N.; Berardi, V.; Linstead, E.; Maoz, U. Examining the utility of nonlinear machine learning approaches versus linear regression for predicting body image outcomes: The U.S. Body Project I. *Body Image* **2022**, *41*, 32–45. [CrossRef] [PubMed]
60. Yates, L.A.; Aandahl, Z.; Richards, S.A.; Brook, B.W. Cross Validation for Model Selection: A Review with Examples from Ecology. *Ecol. Monogr.* **2023**, *93*, e1557. [CrossRef]
61. Chicco, D.; Warrens, M.J.; Jurman, G. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Comput. Sci.* **2021**, *7*, e623. [CrossRef] [PubMed]
62. Rainio, O.; Teuho, J.; Klén, R. Evaluation Metrics and Statistical Tests for Machine Learning. *Sci. Rep.* **2024**, *14*, 6086. [CrossRef] [PubMed]
63. Khoshvaght, H.; Permala, R.R.; Razmjou, A.; Khiadani, M. A Critical Review on Selecting Performance Evaluation Metrics for Supervised Machine Learning Models in Wastewater Quality Prediction. *J. Environ. Chem. Eng.* **2025**, *13*, 119675. [CrossRef]
64. Wisniewski, S.J.; Brannan, G.D. Correlation (Coefficient, Partial, and Spearman Rank) and Regression Analysis. In *StatPearls*; StatPearls Publishing: Treasure Island, FL, USA, 2024.
65. Joseph, V.R. Optimal ratio for data splitting. *Stat. Anal. Data Min. ASA Data Sci. J.* **2022**, *15*, 531–538. [CrossRef]
66. Prokhorenkova, L.; Gusev, G.; Vorobev, A.; Dorogush, A.V.; Gulin, A. CatBoost: Unbiased boosting with categorical features. In *Advances in Neural Information Processing Systems 31*; Curran Associates, Inc.: Red Hook, NY, USA, 2018.
67. Bergstra, J.; Bengio, Y. Random search for hyper-parameter optimization. *J. Mach. Learn. Res.* **2012**, *13*, 281–305.
68. Bischl, B.; Binder, M.; Lang, M.; Pielok, T.; Richter, J.; Coors, S.; Thomas, J.; Ullmann, T.; Becker, M.; Boulesteix, A.-L.; et al. Hyperparameter Optimization: Foundations, Algorithms, Best Practices, and Open Challenges. *WIREs Data Min. Knowl. Discov.* **2023**, *13*, e1484. [CrossRef]
69. Prechelt, L. Automatic early stopping using cross validation: Quantifying the criteria. *Neural Netw.* **1998**, *11*, 761–767. [CrossRef] [PubMed]
70. Miseta, T.; Fodor, A.; Vathy-Fogarassy, Á. Surpassing Early Stopping: A Novel Correlation-Based Stopping Criterion for Neural Networks. *Neurocomputing* **2024**, *567*, 127028. [CrossRef]
71. Chen, H.; Covert, I.C.; Lundberg, S.M.; Lee, S.-I. Algorithms to Estimate Shapley Value Feature Attributions. *Nat. Mach. Intell.* **2023**, *5*, 590–601. [CrossRef]
72. Lundberg, S.M.; Lee, S.-I. A Unified Approach to Interpreting Model Predictions. In *Advances in Neural Information Processing Systems 30*; Curran Associates, Inc.: Red Hook, NY, USA, 2017.
73. Roth, J.P.; Bajorath, J. Machine Learning Models with Distinct Shapley Value Explanations Decouple Feature Attribution and Interpretation for Chemical Compound Predictions. *Cell Rep. Phys. Sci.* **2024**, *5*, 102110. [CrossRef]
74. Shapley, L.S. A Value for n-Person Games. In *Contributions to the Theory of Games II*; Kuhn, H.W., Tucker, A.W., Eds.; Princeton University Press: Princeton, NJ, USA, 1953; pp. 307–317.
75. Hu, N.; Pavlou, P.A.; Zhang, J. On self-selection biases in online product reviews. *MIS Q. Manag. Inf. Syst.* **2017**, *41*, 449–471. [CrossRef]
76. Liu, Q.; Hang, T.; Wu, Y. Unveiling differential impacts of multidimensional urban morphology on heat island effect across local climate zones: Interpretable CatBoost-SHAP machine learning model. *Build. Environ.* **2025**, *270*, 112574. [CrossRef]
77. Mittal, V.; Ross, W.T., Jr.; Baldasare, P.M. The Asymmetric Impact of Negative and Positive Attribute-Level Performance on Overall Satisfaction and Repurchase Intentions. *J. Mark.* **1998**, *62*, 33–47. [CrossRef]
78. Kahneman, D.; Tversky, A. Prospect Theory: An Analysis of Decision under Risk. *Econometrica* **1979**, *47*, 263–291. [CrossRef]
79. Jiang, L.; Diao, X.; Zhang, Y.; Zhang, J.; Li, T. Review of the Charging Safety and Charging Safety Protection of Electric Vehicles. *World Electr. Veh. J.* **2021**, *12*, 184. [CrossRef]
80. Ha, S.; Marchetto, D.J.; Dharur, S.; Asensio, O.I. Topic Classification of Electric Vehicle Consumer Experiences with Transformer-Based Deep Learning. *Patterns* **2021**, *2*, 100195. [CrossRef] [PubMed]
81. Hou, Y.; Wang, P.; Yao, Z.; Zheng, X.; Chen, Z. Enhancing Electric Vehicle Charging Infrastructure Planning with Pre-Trained Language Models and Spatial Analysis: Insights from Beijing User Reviews. *ISPRS Int. J. Geo-Inf.* **2025**, *14*, 325. [CrossRef]
82. Azzouz, L.; Brand, C.; Fawcett, T.; Zhou, Z.; Altaf, M. Beyond the Plug: Enhancing the User Experience at Public Electric Vehicle (EV) Charging Hubs. Insights from a Multi-Site UK Study. *J. Transp. Geogr.* **2026**, *131*, 104530. [CrossRef]
83. Wang, Y.-Y.; Chi, Y.-Y.; Xu, J.-H.; Li, J.-L. Consumer Preferences for Electric Vehicle Charging Infrastructure Based on the Text Mining Method. *Energies* **2021**, *14*, 4598. [CrossRef]
84. Lv, S.; Xiao, A.; Qin, Y.; Xu, Z.; Wang, X. A Decision Framework for Improving the Service Quality of Charging Stations Based on Online Reviews and Evolutionary Game Theory. *Transp. Res. Part A Policy Pract.* **2024**, *187*, 104168. [CrossRef]
85. Ullah, I.; Liu, K.; Yamamoto, T.; Zahid, M.; Jamal, A. Modeling of Machine Learning with SHAP Approach for Electric Vehicle Charging Station Choice Behavior Prediction. *Travel Behav. Soc.* **2023**, *31*, 78–92. [CrossRef]
86. Kuby, M.; Cordova-Cruzatty, A.; Parker, N.C.; King, D.A. EV Charging for Multifamily Housing: Review of Evidence, Methods, Barriers, and Opportunities. *Renew. Sustain. Energy Rev.* **2025**, *210*, 115253. [CrossRef]

87. Fescioglu-Unver, N.; Yıldız Aktaş, M. Electric Vehicle Charging Service Operations: A Review of Machine Learning Applications for Infrastructure Planning, Control, Pricing and Routing. *Renew. Sustain. Energy Rev.* **2023**, *188*, 113873. [CrossRef]
88. Motlagh, S.G.; Li, L. A Review on Electric Vehicle Charging Station Planning: Infrastructure Placement, Sizing, Upgrades, and Uncertainties. *J. Energy Storage* **2026**, *141*, 119325. [CrossRef]
89. Hussain, A.; Islam, M.M.; Shah, S.S.; Zahir, M.H.; Ahammad, A.J.S.; Aziz, M.A. Recent progress in polyetherimide membrane for high-performance lithium batteries. *J. Energy Storage* **2025**, *131*, 117584. [CrossRef]
90. Blonsky, M.; Munankarmi, P.; Balamurugan, S.P. Incorporating Residential Smart Electric Vehicle Charging in Home Energy Management Systems. In *Proceedings of the 2021 IEEE Green Technologies Conference*; IEEE: Denver, CO, USA, 2021; pp. 187–194.
91. Hardman, S.; Jenn, A.; Tal, G.; Axsen, J.; Beard, G.; Daina, N.; Figenbaum, E.; Jakobsson, N.; Jochem, P.; Kinnear, N.; et al. A Review of Consumer Preferences of and Interactions with Electric Vehicle Charging Infrastructure. *Transp. Res. Part D Transp. Environ.* **2018**, *62*, 508–523. [CrossRef]
92. Tversky, A.; Kahneman, D. Loss Aversion in Riskless Choice: A Reference-Dependent Model. *Q. J. Econ.* **1991**, *106*, 1039–1061. [CrossRef]
93. Rempel, D.; Cullen, C.; Bryan, M.M.; Cezar, G.V. Reliability of Open Public Electric Vehicle Direct Current Fast Chargers. *Hum. Factors* **2024**, *66*, 3054–3067. [CrossRef] [PubMed]
94. Bhole, B.; Hanna, B. The Effectiveness of Online Reviews in the Presence of Self-Selection Bias. *Simul. Model. Pract. Theory* **2017**, *77*, 108–123. [CrossRef]
95. Roozkhosh, F.; Yao, X.A. Understanding Electric Vehicle Charging Behavior: A Multidisciplinary Review and Conceptual Framework. *Transp. Res. Part D Transp. Environ.* **2026**, *127*, 105227.
96. Xu, H.; Zhang, D.; Zhang, Y.; Xu, R. HITSZ-HLT at SIGHAN-2024 dimABSA Task: Integrating BERT and LLM for Chinese Dimensional Aspect-Based Sentiment Analysis. In *Proceedings of the 10th SIGHAN Workshop on Chinese Language Processing*; Association for Computational Linguistics: Bangkok, Thailand, 2024; pp. 175–185.
97. Wu, X.; Liao, H.; Ngai, E.W.T. Ranking Hotel Products by Integrating Online Ratings from Multiple Platforms. *Inf. Manag.* **2024**, *61*, 103959. [CrossRef]
98. Qi, J.; Zhang, Z.; Jeon, S.; Zhou, Y. Mining Customer Requirements from Online Reviews: A Product Improvement Perspective. *Inf. Manag.* **2016**, *53*, 951–963. [CrossRef]

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.