

A MATHEMATICAL ORDINAL MULTICLASS FRAMEWORK FOR ELECTRIC-VEHICLE ADOPTION INTENTION BASED ON BAYESIAN-OPTIMIZED XGBOOST

YaJie Liu*, LiMin Hao, ZhangJingYuan Wang

College of Information Engineering, Tarim University, Alaer 843300, Xinjiang, China.

**Corresponding Author: YaJie Liu*

Abstract: As electric-vehicle markets broaden, consumers do not enter with a single, uniform level of adoption intention. This study asks whether Low, Medium, and High intention can be separated reliably in an imbalanced sample and which observed variables do most of the separating. The data contain 50,000 consumers and 21 measures spanning demographics, mobility, charging, cognition, prior experience, and cost. Cleaning, encoding, and standardization produce 25 model inputs. Spearman correlation is used to inspect monotone relationships before a regularized multiclass XGBoost model is fitted with a stratified 80:20 split, stratified five-fold cross-validation, and 100 Bayesian optimization trials. Gain and SHAP are then used to examine overall and class-specific contributions. Mean cross-validation accuracy is 86.88% (SD 0.61%), and accuracy on the independent test set is 87.40%. F1-scores are 0.94 for High, 0.83 for Low, and 0.74 for Medium. The weaker Medium result reflects overlapping distributions and greater within-class variation. Electric-vehicle knowledge and range anxiety are the leading monotone correlates and together account for 0.419 of Gain importance. The findings connect predictive performance with boundary uncertainty and feature contribution in an interpretable ordinal-classification setting.

Keywords: Ordinal multiclass classification; XGBoost; Bayesian optimization; Spearman rank correlation; SHAP; Electric-vehicle adoption

1 INTRODUCTION

Many factors affect the application of electric vehicles, and only one reason cannot be used to explain it. Technology readiness, perceived risk, charging access, household economic conditions and consumers' awareness of electric vehicles may all influence adoption simultaneously, and some of these effects may also act in different ways. With the gradual spread of battery-electric vehicles among more consumers beyond early adopters, a practical problem has arisen: whether these consumers are willing to buy them, and if so, how strong this desire is; can we observe any differences based on available data? Previous machine-learning studies have shown that demographic information, behaviour characteristics and infrastructure conditions can help identify consumers who may be interested in electric vehicles [1]. Consumers' knowledge of electric vehicles is also related to their willingness to purchase [2], and the visibility and convenience of charging facilities may affect how people view the practicability and social acceptability of electric vehicles [3]. Therefore, several related variables should be considered together for adoption intention because there may be non-linear relationships, and a single factor cannot serve as the main explanation.

Another point that needs to be considered is the order of the intention categories. Low, Medium, and High intention are different levels of the same preference, so they are not three completely unrelated categories; however, although this order contains useful information, it does not mean that the distance between two neighboring categories is fixed. In other words, the difference between Low and Medium may not be the same as the difference between Medium and High. There are also two problems in the sample, the number of consumers with High intention is much larger than the numbers in the other two categories, and some predictors provide partly similar information. Because of this, the modeling process has to deal with an ordered outcome, possible nonlinear differences, unequal class sizes, correlation among input variables, and the need to make the model results understandable.

To examine this problem, the study uses 50,000 records and 21 variables, including variables related to demographics, mobility, charging, cognition, experience, and cost. After data processing, the model uses a standardized 25-dimensional input and gives the probabilities of the three ordered intention levels; model performance is then evaluated by using observations that were not included in model fitting. Spearman coefficients are first calculated to examine monotone relationships among the variables and to check whether some predictors contain similar information. After this step, a regularized multiclass XGBoost model is built based on a stratified 80:20 split, while stratified five-fold validation and 100 Bayesian tuning trials are also used during model development. Global Gain importance and class-specific SHAP values are then used to examine the role of different predictors and how they are related to the probabilities produced by the model.

The three components of this paper are as follows. First, adoption intention is divided into three ordered categories, and thus, instead of reducing the outcome to just a yes-or-no purchase decision, all differences among the various levels of intention can be preserved. Second, rank screening, regularised boosting, Bayesian tuning, stratified validation and additive explanation are added to the same analysis process, and thus the main modelling steps are relatively complete.

Third, model performance is evaluated for both the entire sample and each intention category; thus, the weaker classification result for medium intention can be linked to the overlap of observed features, rather than being explained solely as a model error. Therefore, the above analysis can provide some quantitative evidence for differences in the willingness of consumers to adopt, as well as how cognitive, infrastructural and behavioural factors affect the classification results.

2 DATA STRUCTURE AND MATHEMATICAL PREPROCESSING

2.1 Sample Space and Response Distribution

The data are 50,000 observations of 21 predictors: age, annual income, education, city type, commuting and weekly travel, current vehicle type and age, fuel expense, public and home charging access, electricity cost, environmental and policy awareness, technology affinity, range anxiety, concern about battery replacement, electric-vehicle knowledge, previous experience, monthly energy use, and monthly charging cost. Adoption intention is divided into three levels: High, Medium and Low. These categories account for 59.3%, 24.2% and 16.5% of the sample respectively; thus, the largest group is about 3.6 times as large as the smallest.

No duplicate records were found in the duplicate check. Missing numerical values are filled with the median, and missing categorical values are filled with the mode. Cap the outliers of continuous variables using the interquartile range method. Education is mapped to 0-3, city type is mapped to 0-2, and current vehicle type is represented by one-hot encoding. The above transformations generate 25 model inputs without imposing an artificial Euclidean order on nominal vehicle categories.

2.2 Standardization and Redundancy Control

All numerical predictors are standardized using the training-sample mean and standard deviation. For a raw value, the standardized coordinate is

$$Z = \frac{x - \mu}{\sigma} \quad (1)$$

so each transformed feature has zero mean and unit variance in the fitted sample. Rank and linear correlations reveal two strong redundancies: monthly charging cost with monthly energy consumption (0.82), and weekly travel distance with monthly energy consumption (0.94). One variable from each highly redundant pair is retained when necessary to limit repeated information.

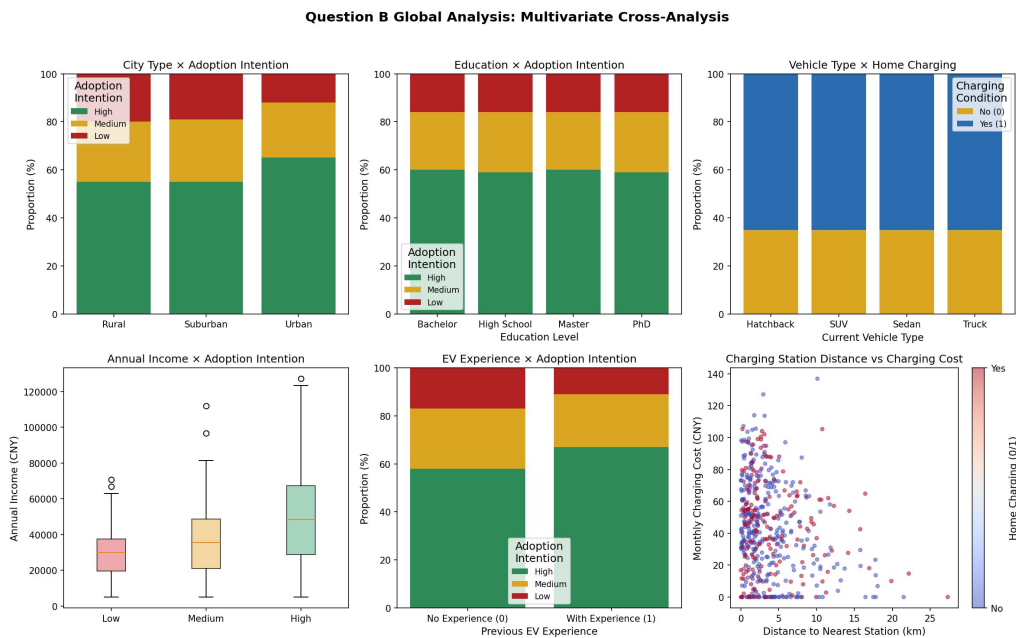


Figure 1 Cross-Sectional Distributions and Grouped Comparisons in the Analyzed Data

Cross-sectional distributions and grouped comparisons in the analyzed data is shown in Figure 1.

3 RANK-BASED FEATURE ANALYSIS AND PROBLEM FORMULATION

3.1 Spearman Association

The result is in an ordered form, several predictors do not follow a normal distribution, and thus Spearman's rank coefficient is used to describe the monotonic relationship between variables. The coefficient for paired ranks with differences is

$$\rho = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2-1)} \quad (2)$$

Electric-vehicle knowledge shows the strongest positive association, $\rho = 0.38$, while range anxiety shows the strongest negative association, $\rho = -0.35$. Technology affinity and environmental awareness have coefficients of 0.28 and 0.25, while income and education show weaker relationships with adoption intention, with values of 0.12 and 0.08. These results suggest that a nonlinear model may be more suitable, especially for the cognitive variables; at the same time, travel and infrastructure variables are still kept in the model, so possible interactions among different variables can also be considered. Figure 2 shows the ranked relationships together with the wider feature profile.

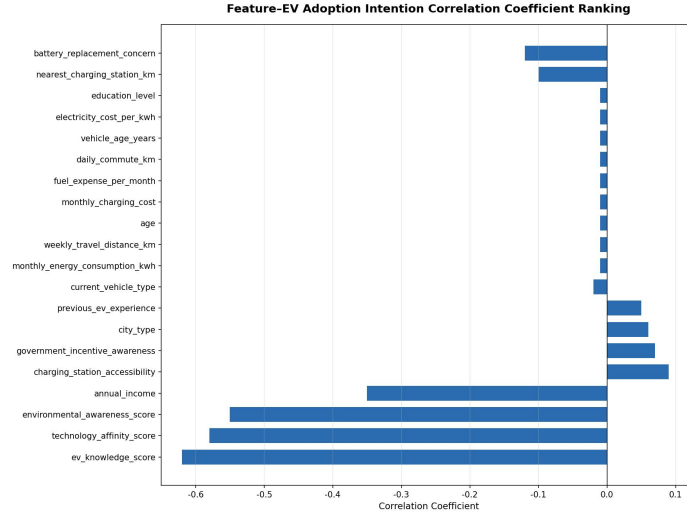


Figure 2 Spearman Ranking and Multidimensional Feature Profile for Adoption Intention

3.2 Ordinal Multiclass Mapping

Although order is used in data interpretation, the predictive stage estimates all three posterior probabilities simultaneously to avoid forcing equal distances between Low, Medium, and High. If the boosted score for class k is denoted by the corresponding class function, the Softmax mapping is

$$p(y_i = k | x_i) = \frac{\exp(f_k(x_i))}{\sum_{j=1}^K \exp(f_j(x_i))} \quad (3)$$

where the number of classes is three. The predicted label is the class with the largest posterior probability. This probability formulation supports both global accuracy and class-specific error analysis.

4 REGULARIZED GRADIENT-BOOSTING MODEL

4.1 Additive Objective

XGBoost builds an additive ensemble of classification trees. At boosting iteration t , a new tree is selected by minimizing the sum of multiclass loss and model complexity:

$$L^{(t)} = \sum_{i=1}^n l(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)) + \sum_{k=1}^K \Omega(f_k) \quad (4)$$

The loss is multiclass cross-entropy. Tree complexity is controlled by

$$\Omega(f) = \gamma T + \frac{1}{2} \lambda \|w\|^2 \quad (5)$$

where T is the number of leaves, w is the leaf-weight vector, and the two nonnegative coefficients penalize tree size and weight magnitude. The regularizer reduces variance and discourages unnecessarily complex partitions of the 25-dimensional space.

4.2 Bayesian Hyperparameter Optimization

The training set contains 40,000 observations and the held-out test set contains 10,000. Within the training set, stratified five-fold cross-validation preserves the three class proportions. Optuna performs 100 Bayesian trials and maximizes validation accuracy. The search ranges are 100–1000 trees, depth 3–12, learning rate 0.01–0.30, row and feature subsampling 0.60–1.00, L1 and L2 coefficients 0–10, and minimum child weight 1–10. Early stopping terminates training after 50 rounds without validation improvement.

The parameter search is used to balance model fit and regularisation. Deeper trees and a larger ensemble can learn more intricate patterns, but shrinkage, subsampling and penalty terms limit model complexity. Stratify the five validation folds to prevent the large High group from having too much impact on the evaluation of unseen data.

4.3 Evaluation Functionals

Overall accuracy records the proportion of labels predicted correctly,

$$\text{Accuracy} = \frac{\sum_{i=1}^N I(y_i = \hat{y}_i)}{N} \quad (6)$$

Precision, recall, and F1 are also calculated separately for each class. That class-level view is necessary because a high aggregate accuracy can coexist with poor separation of the intermediate category.

5 PREDICTIVE PERFORMANCE AND ERROR GEOMETRY

5.1 Cross-Validation and Test Results

The optimised model has an average accuracy of 86.88% for the stratified folds and a standard deviation of 0.61%. The results of the different folds are relatively close, so the model is not very sensitive to a single split. An independent test set of 10,000 records achieved an accuracy of 87.40%, which was approximately equal to that in the validation (Table 1).

Table 1 Performance of the Optimized Model for the Three Intention Classes

Class	Precision	Recall	F1-score	Support
High	0.94	0.94	0.94	5,934
Low	0.84	0.82	0.83	1,650
Medium	0.74	0.74	0.74	2,416
Weighted mean	0.87	0.87	0.87	10,000

The three categories have different results. High Intention has a relatively high classification accuracy, partly because of the large sample size and distinct feature pattern. Low intention is also classified reasonably well despite having fewer observations. Medium intention has the lowest F1-score because its cognitive and behavioural measures overlap more with the other two classes and show greater variation within the group. The confusion matrices in Figure 3 show the same pattern and most of the errors are Medium intentions.

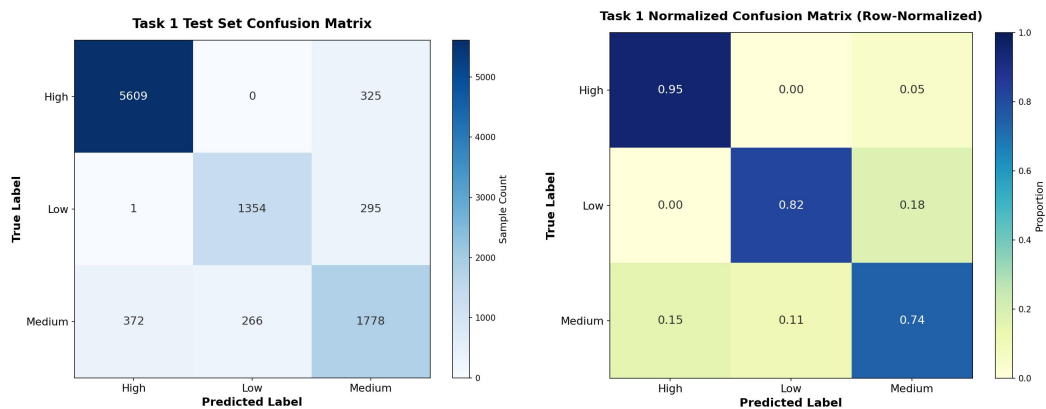


Figure 3 Confusion Matrices for the Independent Test Set

5.2 Boundary Errors

The confusion matrix shows almost no direct error from Low to Medium. Most errors occur between Medium and High or between Low and High. Of the relevant observations, 5.6% of the Low group are classified as High; among High observations, 6.2% are classified as Low and 4.4% as Medium. The misclassified cases do not have one common pattern. Some respondents have favorable cognitive measures but weaker supporting conditions, while some in the High group still report clear concerns about range or batteries.

Mathematically, these cases indicate overlap in class-conditional feature distributions rather than unstable optimization. A future cost-sensitive objective or minority-aware resampling scheme could increase the penalty assigned to intermediate-class errors, but no unreported performance gain is assumed here.

6 INTERPRETABLE FEATURE CONTRIBUTION

6.1 Gain Importance

Tree-based gain measures the average reduction in split impurity attributable to a feature. Across all trees, the importance of feature j is summarized by

$$\text{Gain}_j = \frac{1}{M} \sum_{m=1}^M \Delta I_{j,m} \quad (7)$$

where M is the number of relevant split occurrences and the impurity reduction is evaluated at each occurrence. The ranking is dominated by cognitive variables. Six of the ten leading predictors belong to the cognition domain and together account for more than 71% of the reported importance. Electric-vehicle knowledge has importance 0.223 and range anxiety 0.196, giving a combined contribution of 0.419. Figure 4 presents the corresponding gain-based ranking and grouped infrastructure and energy-consumption distributions, allowing the dominant cognitive predictors to be compared with secondary contextual variables.

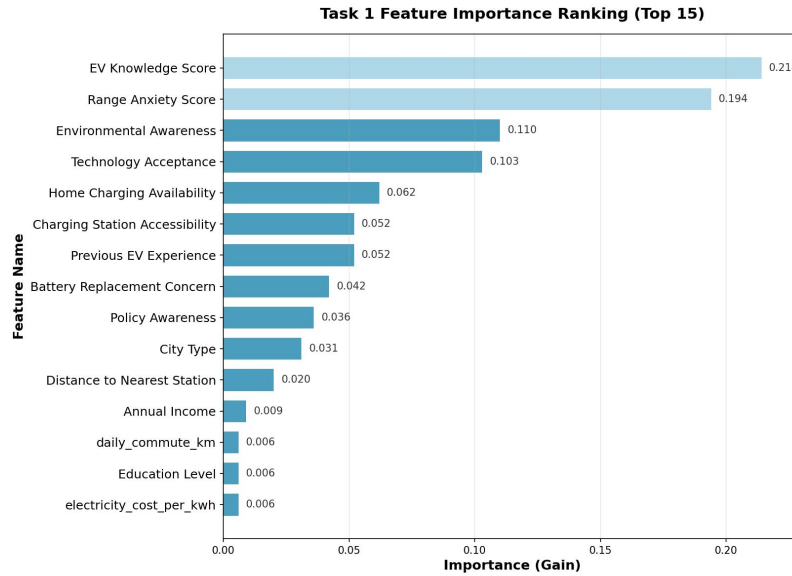


Figure 4 Gain-Based Feature Ranking and Grouped Infrastructure and Energy-Consumption Distributions

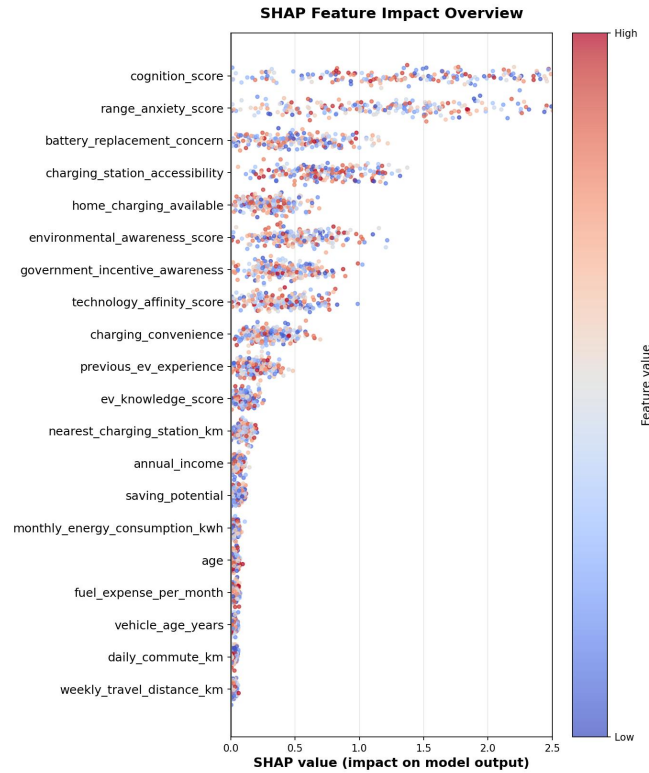


Figure 5 Class-Specific SHAP Results by Intention Level

6.2 Additive SHAP Interpretation

Gain describes global split utility, whereas SHAP values decompose each individual prediction into additive feature effects. According to the SHAP results, electric-vehicle knowledge, range anxiety, environmental awareness and

technology affinity are frequently among the variables with relatively high contributions. Generally speaking, a higher level of knowledge and environmental consciousness will increase the predicted likelihood of a favourable attitude; otherwise, it is reduced by range anxiety. The above patterns describe the fitted model and should not be interpreted as causal effects. Figure 5 shows the SHAP results for low, medium and high intentions separately, making it easier to compare the differences in direction and magnitude.

7 MATHEMATICAL DISCUSSION

7.1 Model Interpretation

The fitted model can be considered as a piecewise-constant approximation of the probability surface for the three classes. Rank screening mainly shows monotone patterns for individual variables, and boosted trees can capture interactions that pairwise coefficients may miss. Softmax changes the three fitted scores into probabilities for the three classes that sum to one. Regularization limits the number and size of leaves, and repeated stratified validation is used during Bayesian optimisation to control the balance between bias and variance.

Charging variables have a relatively small weight compared to the main cognitive variables, but they still contribute to the model. The importance of home charging availability is 0.057, and public charging access has an importance of 0.051. The higher value for home charging is in line with evidence that private access may be more closely related to purchase intention than the density of public stations [4]. Other published results have also shown that the relationship between vehicle range and charging infrastructure varies in different study areas [5].

The general result is consistent with previous studies that have linked adoption intention to technology familiarity, perceived risk, environmental concern and charging conditions [6-8]. Recent studies on explainable learning have also used boosted trees and additive attribution to investigate differences in adoption behaviour [9-10]. Cognitive measures account for most of the model's gain importance in this sample; demographic variables contribute relatively little to additional differentiation.

7.2 Uncertainty, Limitations, and Extension

Medium Intention is relatively uncertain. For several variables with monotonic trends, this group is located near the middle and thus has a small distance to its neighbours. Its F1-score of 0.74 shows that the two distributions have some overlap, unequal class sizes, and may be subjectively labelled. Cross-sectional data are used in this study, so causality cannot be established. The distribution of this group will change in the future due to changes in technology, charging access and consumer knowledge.

A stacking model is also employed for stability verification. Logistic Regression and XGBoost are used to make predictions, and then a logistic meta-learner combines these predictions; this approach can reduce errors caused by linear and non-linear models. In the future, compare ordinal loss functions, cost-sensitive boosting and class-weighted cross-entropy under the same holdout settings.

8 CONCLUSION

This paper is an ordinal multi-class model of electric vehicle adoption intention. Median or mode imputation, IQR capping, categorical encoding and Z-score scaling are applied to the 21 observed variables to create 25 model inputs. Spearman coefficients show that knowledge and range anxiety have the strongest monotonic relationships with intention. After Bayesian tuning, multiclass XGBoost reaches $86.88\% \pm 0.61\%$ accuracy in cross-validation and 87.40% on the test set. At the class level, the size of the sample does not determine how difficult classification is; Medium intention has many more errors than Low, even though it is relatively frequent. Gain and SHAP also show that cognitive measures contribute most to the fitted model. Results of both classification performance and information on class separation, feature contribution, and remaining uncertainty are shown.

Several deficiencies should be pointed out. The data are cross-sectional; therefore, the estimates only indicate prediction and cannot establish causality. Class proportions and feature effects may change with battery technology, charging network construction, incentives and consumer familiarity, etc. The three intention levels also simplify a continuous preference, and the lower F1-score for Medium shows that overlapping profiles are still difficult to separate. Future studies can test the model with samples from different times and regions, compare ordinal losses with cost-sensitive and class-weighted objectives, and examine changes in charging access, vehicle experience and policy exposure. External validation and calibration should also be repeated to determine whether the fitted boundaries and feature contributions remain consistent across markets.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

- [1] Bas J, Cirillo C, Cherchi E. Classification of potential electric vehicle purchasers: A machine learning approach. *Technological Forecasting and Social Change*, 2021, 168: 120759.
- [2] Jaiswal D, Kant R, Singh P K, et al. Investigating the role of electric vehicle knowledge in consumer adoption: evidence from an emerging market. *Benchmarking: An International Journal*, 2021, 29(3): 1027-1045.
- [3] White L V, Carrel A L, Shi W, et al. Why are charging stations associated with electric vehicle adoption? Untangling effects in three United States metropolitan areas. *Energy Research & Social Science*, 2022, 89: 102663.
- [4] Kristoffersson I, Pyddoke R, Kristofersson F, et al. Access to charging infrastructure and the propensity to buy an electric car. *Transportation Research Part D: Transport and Environment*, 2025, 139: 104588.
- [5] Rosales-Tristancho A, Carazo A F, Brey R, et al. Meta-analysis of consumers' trade-offs between range and charging infrastructure for electric vehicles. *Transportation Research Part D: Transport and Environment*, 2024, 134: 104342.
- [6] Pamidimukkala A, Kermanshachi S, Rosenberger J M, et al. Barriers and motivators to the adoption of electric vehicles: A global review. *Green Energy and Intelligent Transportation*, 2024, 3(2): 100153.
- [7] Pandita D, Bhatt V, Kumar V V R, et al. Electrifying the future: analysing the determinants of electric vehicle adoption. *International Journal of Energy Sector Management*, 2024, 18(6): 1767-1786.
- [8] Fu X. Understanding the adoption intention for electric vehicles: The role of hedonic-utilitarian values. *Energy*, 2024, 301: 131703.
- [9] Pamidimukkala A, Kermanshachi S, Rosenberger J M, et al. Adoption of electric vehicles: An empirical study of consumers' intentions. *Transport Economics and Management*, 2024, 2: 359-366.
- [10] Kisting D, Titz M, Bottcher P C, et al. Revealing drivers of green technology adoption through explainable artificial intelligence. *Advances in Applied Energy*, 2025, 20: 100242.