Travel Behaviour and Society 31 (2023) 78–92

Contents lists available at ScienceDirect

Travel Behaviour and Society

journal homepage: www.elsevier.com/locate/tbs

Modeling of machine learning with SHAP approach for electric vehicle charging station choice behavior prediction

Irfan Ullah [a,b], Kai Liu [a,*], Toshiyuki Yamamoto [c], Muhammad Zahid [d], Arshad Jamal [e]

[a] School of Transportation and Logistics, Dalian University of Technology, Dalian 116024, China
[b] Department of Business and Administration, ILMA University, Pakistan
[c] Institute of Materials and Systems for Sustainability, Nagoya University, Nagoya 464-8603, Japan
[d] Department of Civil, Geological and Mining Engineering, Polytechnique Montreal, Montreal, Canada
[e] Transportation and Traffic Engineering Department, College of Engineering, Imam Abdulrahman Bin Faisal University, Dammam, Saudi Arabia

ARTICLE INFO

Keywords:
Electric vehicles (EVs)
Charging choice behavior
Machine learning (ML) algorithm
SHAP
Real-world data

ABSTRACT

Growing electric mobility makes it difficult for electric vehicles (EVs) to charge adequately while charging infrastructure capacities are limited. Due to the prolonged charging times, precise planning is needed, which necessitates knowing the availability of charging stations. In addition, inconsistencies in charging facilities and illogical charging arrangements cause partial queuing and idling of charging stations. To tackle these issues, it is necessary to first understand EV charging station choice behavior and its influence. This study examines EV charging station choice behavior and aims to find the best prediction method. This study implements a novel interpretable machine learning (ML) framework to predict EVs' charging station choice behavior. The experiment was based on two years of real-world normal and fast charging event data from 500 EVs in Japan. The results revealed that the XGBoost model achieved the highest accuracy compared to the other ML classifiers in predicting charging station choice behavior. Furthermore, this study employed the newly developed SHAP approach to identify feature importance and the complex nonlinear and interactive effects of various attributes on charging station choice behavior. This study suggests that combining ML models with SHAP has the potential to develop an interpretable ML model for predicting EV charging station choice behavior.

1. Introduction

Electric vehicles (EVs) are currently regarded as one of the essential measures to lessen transportation's negative environmental impact. EVs have become an essential component of the automotive sector. In 2019, EV sales hit 2.1 million, following a 40 % annual growth trend (Shibl et al., 2021). Moreover, EV charging stations have become an essential part of global infrastructure, with 7.3 million charging stations installed globally in 2019 and a 60 % increase compared to 2018 (Shibl et al., 2021). Furthermore, global EV sales are predicted to reach 43 million by 2030, accounting for 30 % of all vehicles attributable to the advancement of fast charging technologies (ElMenshawy and Massoud, 2020; ElMenshawy and Massoud, 2020). However, despite recent technological advancements, they are still constrained by their battery capacity. Their batteries are heavy, large, and expensive, resulting in a limited driving range (Westin et al., 2018; Ullah et al., 2021). Moreover, charging infrastructure is limited, and charging EVs takes a longer time than refuelling diesel vehicles (Liu et al., 2021; Liu et al., 2022). These main drawbacks impede EV adoption. Because EVs have limited driving ranges and prolonged charging times duration, policymakers gave top priority to the development of charging stations to foster the expansion of the EV market (Pan et al., 2019; van Heuveln et al., 2021). To further encourage the use of EVs in the future, a problem that needs to be solved is how to distribute charging stations with limited resources, particularly to handle urban and rural growth. The driving range of EVs is limited in comparison to those powered by internal combustion engines (ICEVs). Therefore, EV drivers must base their trip planning on the availability of chargers (Chakraborty et al., 2022). These problems make it challenging for motorists to select EVs as a transportation mode. Therefore, more caution at charging stations and route planning is needed, particularly for commercial vehicles, because different vehicles have different waiting and charging times. As a result, it is critical to

* Corresponding author.
E-mail addresses: irfanktk@mail.dlut.edu.cn (I. Ullah), liukai@dlut.edu.cn (K. Liu), yamamoto@civil.nagoya-uac.jp (T. Yamamoto), zahid7984@gmail.com (M. Zahid), ajjamal@iau.edu.sa (A. Jamal).

https://doi.org/10.1016/j.tbs.2022.11.006
Received 21 April 2022; Received in revised form 5 October 2022; Accepted 19 November 2022
Available online 30 November 2022

investigate how charging behavior is influenced by charging stations using real-world charging event data, which have rarely been included in earlier studies.

Currently, for EVs, two types of electric power supply methods are available: battery replacement and charging (Wang et al., 2011). Their charging magnitude is usually classified into three levels (1, 2, and 3). Level-1 charging is the utmost time-consuming mode among all available charging modes. In the US, Level-1 charging is a standard 120 V/15 A. It takes 6–12 h for a PHEV and 12–20 h to fully charge EVs. The total cost of the Level-1 charging station is expected to be approximately $500-$880 (De Sousa et al., 2010; San Román et al., 2011). However, it requires special equipment and connected devices for public or home charging stations (Rawson and Kateley, 1999). Level-2 charging is the basic technology that dedicates basic facilities to public and private facilities. Currently, Level-2 charging is used as a standard 240 V/80 A. It takes 4–8 h for a PHEV and 6–14 h to fully charge EVs. EV owners choose Level-2 charging because of its shorter charging time and standard vehicle charger connectors. The total cost of a Level-2 charging station is approximately 2150$ (San Román et al., 2011). For example, an additional $3,000 costs have been levied by the Tesla Roadster charging network (Berdichevsky et al., 2006). Level-3 charging is connected to the city, plus a highway gas station equivalent to a petrol station. It usually works in a three-phase circuit of 480 V or higher (Rahman et al., 2016) and needs an off-board connection to provide a controlled AC-DC connector. The Level-3 charging station is much lower in the residential area. It takes 15 min to 3 h for a PHEV and 1 to 4 h to fully charge EVs. However, the setup costs of the Level-3 charging station are between $30,000 and $160,000 (Thomas, 2009). Furthermore, it seems that current charging stations can support different types of charging demand. A different connection can be used for different charging demands; for example, SAE J1772 is the industry standard for all EVs performing Level-1 and Level-2 charging. SAE J1772 can connect to either the cord that came with the EV or the Level 2 charger available outside. The connector CHAdeMO was introduced for DC fast chargers. It was originally intended to be the industry standard, having been developed through the collaboration of five major Japanese automakers. The connector CSS, shortly after the CHAdeMO was introduced, a second connector known as the combined charging system (CCS) was implemented. CCS connectors differ from CHAdeMO in that they support both AC and DC charging on the same port. To charge at Level 1 or 2, CHAdeMO-equipped vehicles require an additional J1772 connector cord. The Tesla supercharger always opted to pave its own path in the EV market, and that is no different from its Supercharger connector. This unique connector is found on all Tesla models in North America, although it does offer CCS and CHAdeMO adapters in several markets.

Levels 1 and 2 generally refer to slow or normal charging modes, while level 3 is known as a fast charging mode. The most popular charging modes are normal, fast, and wireless charging. Normal charging is more appropriate in places where people spend more time, such as offices during business hours and residential areas throughout the night (Liao et al., 2017). As a result, they provide significant flexibility and convenience for charging batteries while EVs are parked overnight at home or all day at work. Simultaneously, charging during travel is frequently possible due to a low state of charge (SOC) at the start of the journey or because the destination is too far. Due to the urgency of the charge, fast charging is usually necessary in this case. Despite their excellent charging efficiency, fast charging stations are limited to public infrastructure because of high voltage requirements. Understanding how incentives affect EV adoption and how to install charging infrastructure necessitates the development of a concrete behavioral prediction model that can capture EV users' charging patterns. This effect is closely linked to the utilization of charging infrastructure and, as a result, EV users' charging station choice behavior.

Charging behavior is influenced by several factors. Despite the importance of charging behavior modeling, few studies have examined it due to a lack of actual revealed preference data. Earlier studies examined EVs' charging behavior from a different aspect. Charging behavior studies are classified into two categories: enroute charging (fast charging) and after-trip charging (normal charging). Enroute charging usually includes both charging and route planning, particularly when an EV is unable to complete the entire trip. Drivers need to charge EVs and frequently detour during the journey. Sun et al. investigated fast charging station selection during a trip. The finding shows that private users could usually detour up to 1,750 and 750 m on working and nonworking days, whereas commercial users could only detour 500 m on both working and nonworking days (Sun et al., 2016). Yang et al. suggested a model to analyze EV charging behaviors (Yang et al., 2016). Meanwhile, it has been exposed that the initial SOC of the EV at the origin is the important factor for charging decisions, and at the destination, it has become an essential factor for EV drivers' route choice. Jabeen et al. proposed a study investigating EV charging behavior while accounting for charging time and cost. The findings show that the majority of drivers charged their EVs at home. When drivers charge their EVs at public charging stations, they are worried about the charging duration time. A study examined the main influencing factors of EV drivers' charging behavior (Azadfar et al., 2015). The results revealed that the charging infrastructure facilities, charging cost, and performance of the battery were significant factors for charging choice behavior. The authors conducted a study of 79 EV drivers to examine the impact of the driver's psychological state on charging behavior (Franke and Krems, 2013). The results suggested that EV driver charging behavior is influenced by EV driver knowledge, the suitable pricing range, and EV usage efficiency. Adornato et al. conducted a survey of EV drivers in Michigan, USA, to find the EV's charging location and estimate its charging time. Furthermore, they proposed a model to predict charging demand and energy consumption (Adornato et al., 2009). The results showed that drivers frequently charged EVs at shopping centers, at work and at home with corresponding charging times of 30 min, 9.4 h, and 3.8 h, respectively. Most research has focused on applying statistical techniques to analyze EV charging behavior. These research studies investigated the statistical regularity of charging choice behavior. Furthermore, based on real-world EV running data, scant literature has examined charging station behavior.

Several studies have developed supervised machine learning (ML) algorithms for EV charging behavior prediction. Frendo et al. employed a simple regression model to predict EV departure times (Frendo et al., 2020). The prediction model was trained using historical data spanning three years and included over 100,000 charging sessions. The results revealed that the XGBoost model outperformed the other models. Xiong et al. employed a linear regression (LR) model to predict EV user behavior regarding session duration and start time (Xiong et al., 2017). However, EV charging behavior predictions were incorporated into the smart charging algorithm to provide grid stabilization. Xu established a support vector machine (SVM) to predict EV arrival and departure times. Three years of EV commuter charging data were used in this study. The result shows that the average MAPE values are 2.85 % (arrival time) and 3.7 % (departure times). The suggested model outperformed the simple persistence reference forecasting model (Xu, 2017). Chung et al. utilized several standard ML algorithms, random forest (RF), k-nearest neighbor (KNN), and SVM, to predict energy consumption and session duration from two types of charging data sets (Chung et al., 2019). The first data set contained charging sessions for nonresidential charging behavior, while the second data set was obtained from residential charging behavior. The findings show that SVM is superior (SMAPE 10.54 %) to other algorithms for session duration; for energy consumption, RF performed well (SMAPE 8.73 %). Songpu et al. adopted several widely used ML algorithms, RF, adaptive boosting (AdaBoost), gradient boosting (GBoost), and naive Bayes (NB), to predict household day-head EV charging occurrence time and the "no charge" at day (Ai et al., 2018). Furthermore, in this study, the proposed algorithms were combined into an ensemble algorithm to achieve better results. Majidpour et al. proposed various ML algorithms, SVM, RF, and modified pattern-based

Table 1
Raw data of a particular EV.

GPS Location	Vehicle ID	Vehicle type	Start SOC (%)	End SOC (%)	Arrival time	Departure time	Charging duration/min	Charging mode
***	3001974	Commercial	97	100	2/5/2011 11:36	2/5/2011 12:09	33	Normal charging
***	3001969	Commercial	30.5	76.5	3/3/2011 21:11	3/3/2011 21:37	26	Fast charging
***	3001969	Commercial	52	80	4/6/2011 19:37	4/6/2011 19:53	16	Fast charging
***	3002152	Private	62	100	9/26/2011 23:21	9/27/2011 1:06	105	Normal charging
***	3002152	Private	59	84.5	9/27/2011 13:14	9/27/2011 13:35	21	Fast charging
***	3002152	Private	57.5	100	9/27/2011 23:38	9/28/2011 1:38	120	Normal charging

The GPS location information is replaced by "***" due to privacy protection.

Table 2
Numbers of fast and normal charging events.

Charging events	Commercial vehicles	Private vehicles	Total
Fast charging	2606	6526	9132
Normal charging	52746	38678	91424

sequence forecasting (MPSF), to predict the charging outlet and energy consumption (Majidpour et al., 2014). The result suggested that MPSF having (SMAPE 14.06 %) is outperformed. Yang et al. utilized a binary logistic regression model to classify whether the driver will charge the vehicle at a fast charging station (Yang et al., 2020). The data were collected from 130 private EVs in Beijing containing 15,752 trajectories; among 2161 have fast charging station behavior. The suggested approach yields the highest accuracy (0.894) compared to multivariate linear regression (MLR) and univariate linear regression (ULR). Frendo et al. employed XGBoost and ANN to predict charge profiles in the workplace (Frendo et al., 2020). The training data set consists of charging operations from a heterogeneous EV fleet. The findings show that XGBoost yields a good prediction with MAE 126 W and MAE 0.06. Lu et al. proposed an RF algorithm to predict daily charging times and charging capacity. To achieve a short-term forecast, this method was accomplished with a classification and regression tree algorithm (Lu et al., 2018). The current data's application form in the algorithm is established, and the prediction algorithm's accuracy is proven reliable and practical.

In summary, it is essential to examine EV charging station choice behavior and determine which factors are more important for normal and fast charging stations. Understanding and incorporating the drivers' behavior and preferences is extremely important for the selection and optimal deployment of EV charging infrastructure. For example, it can infer the demand/priority for a specific charging network/type under different scenarios, such as existing SOC, type of vehicle, and temporal and environmental conditions. Furthermore, the modeling of EV charging choice behavior has undergone several stages of transformation. Initially, researchers focused on statistical modeling (discrete choice models) represented by logistics families. However, these models frequently fail to capture the nonlinear relationship between explanatory and response variables. Other main objections to using these approaches in behavior research are unobserved heterogeneity and endogeneity. To overcome these issues, researchers developed ML models for better predictive ability; they criticized the implication behind the black box. To overcome the drawback of the 'black box' and lack of output interpretability, researchers used simple fractures important analysis to highlight the relative significance of each independent variable on the response variable. Now, a reflection exists on the traditional ML modeling framework and discrete choice theory, returning to the pristine but strong charging behavior theoretic support. However, the application of ensemble tree-based ML models with high accuracy and good interpretability in the fields of charging station choice behavior remains scarce, and further research is needed.

Based on the above discussion regarding EV charging station choice behavior prediction, this study developed a novel framework to understand charging station choice behavior by leveraging an interpretable ML approach, aiming to address the research gaps in this area. Specifically, we need to comprehend how each predictor affects the prediction results. To the best of our knowledge, this is one of the first applications in this area of the ML model with SHAP analysis to the field of EV charging station choice behavior. Unlike previous shallow structure approaches, it is unclear how the features affect prediction outcomes. This study disentangles how various explanatory variables influence charging station choice behavior using the SHAP approach to decipher the nonlinear and interactive effects. Therefore, the aim of this study is to develop an interpretable prediction model for EV charging station choice behavior and its influencing factors. Hence, this research identifies the key factors and estimates their effective ranges and thresholds impact on charging choice. We believe that employing this novel technique to provide interpretation of these factors is a qualitative leap. This study provides valuable insight to a researcher who can benefit from the high predictability of nonparametric ML models while addressing related empirical issues and comprehending how to interpret the modeling outcomes.

The rest of the paper is organized as follows. Section 2 describes the proposed research methodology, including the data collection, preprocessing, and prediction methods. Section 3 provides detailed results and analysis. Section 4 provides a sensitivity analysis. Section 5 presents the significance and comparison of employed ML models and finally, Section 6 concludes the paper.

2. Research methodology

2.1. Data collection and preprocessing

In this section, we first describe the data collection of EV charging station events. The Japan Automobile Research Institute collected probe data of 500 EVs used by commercial and private vehicles. Commercial vehicles are fleet vehicles, including business and government vehicles, and private vehicles are household vehicles. The detailed information provided by the probe included odometer reading, clock time, location (latitude and longitude), vehicle state (driving, fast charging or normal charging), heater, A/C on/off, and battery start and end SOC. However, the data trial time was approximately-two years; the sample of private vehicles was done within the last 12 months. Not all vehicles took part in the trial at the same time. However, this trial does not reveal the driver's demographic and socioeconomic characteristics; it only provides regional information where each vehicle is registered, as well as the type of business in the case of commercial vehicles. Table 1 shows the raw data of a particular EV, such as vehicle type, ID, start SOC, end SOC, arrival time, departure time, and charging mode. The raw event-driven data are produced directly from the raw probe EV. The raw probe EV data are prone to errors at both the collecting and receiving ends. For instance, traveling events with a start SOC less than 0 % and an end SOC larger than 100 % make up 0.5 % of the data. 7% of the data are less than 2 min of charging time, or SOC variations between the start and end of the event are less than 1 %. As a result, it is therefore essential to remove these error-prone event-driven data. After removing error-prone event-driven data, the final data were acquired from 239 private vehicles with 6526 fast charging events, 38,678 normal charging events, 247 commercial vehicles with 2606 fast charging events, and 52,746 normal charging events. The detailed events are given in Table 2.

2.2. Prediction methods

ML is a field of study that has developed numerous induction methods that train machines to examine data and extract it. We aim to predict the charging station choice behavior, and the ML algorithm for classification or classifiers is the best fit. The ML classifier assigns a label or a (class) to a new unknown observation. Predicting the class, the ML classifier is trained on a set of previously observed data, i.e., samples with known input features and output class. The input features are usually nominal, ordinal, or numerical. The background details of the supervised ML classifier are included in this section.

2.2.1. Logistic regression (LR)

LR is a supervised learning approach based on a generalized linear regression analysis model. It is a mathematical model that estimates the probability of belonging to a specific class. LR is a data mining technique frequently used to predict classification problems. Finding a prediction function, generating a loss function, and determining regression parameters that minimize a loss function are the three steps in implementing LR. LR analyzes a data set with one or more independent variables. The approach predicts a discrete output from a set of variables that can be discrete, binary, or a combination of any of these (Weng and Yang, 2015). The primary goal of LR is to determine the relationship between the dependent and explanatory variables. When the dependent variable is nominal, two types of LR techniques can be used: binary and multinomial. The binary dependent variable is divided into two categories, commonly written as 0 and 1. For this study, two levels of charging stations (the response variable) were considered. In this regard, the binary LR approach can estimate the likelihood of each possible outcome. Two outcomes were encoded into the target variable. The LR modeling function shows the link between the likelihood of a specific class and the set of independent variables. Equation 1 is a common LR model equation for the current problem. The parameters β_0 and β_1 are identified via iteration, and the ε error term is distributed by the standard logistic function (Walsh et al., 2019).

$$y(x) = \begin{cases} 1, \beta_0 + \beta_1 x + \varepsilon \\ 0, Otherwis \end{cases} \tag{1}$$

The logistic function is used to give $y(x)$, the approximate model as a probability of the response variable being one of two binary values (normal and fast charging).

$$y(x) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x)}} \tag{2}$$

2.2.2. Naive Bayes (NB)

The NB classifier is a probabilistic ML algorithm used to solve classification problems. It is a simple method for modeling classifiers that supposes the observed features are conditionally independent. Bayesian classifiers employ conditional probabilities to categorize response variables based on input variables. It calculates class membership probability, such as the likelihood that a given tuple belongs to a specific class. NB classifiers also model excellent accuracy and speed when applied to an extensive and large database (Zhu et al., 2021; Zhao et al., 2020). It simply assumes that the attribute value does not affect the conditional independence of a given class. It is most effective when the input is high. NB has a strong mathematical base and consistently high classification efficacy. NB excels in small-scale data, handles several classification tasks and is well suited for incremental training (Youn and Jeong, 2009). It generally outperforms more advanced algorithms and is particularly effective in everyday applications such as document classification and spam detection. In this study, the charging station data set is first trained, and then a prediction model is developed. Predictions can be made regarding required conditions, as well as the criticality of the charging station. It is based on probability and Bayesian decision theory. This approach is described as choosing the decision with the highest probability. Using probability and conditional probability, the approach attempts to determine the class to which the relevant input belongs. As a result, it employs the following Bayesian rule (Harrington, 2012).

$$p(c_i|\mathrm{x},\mathrm{y}) = \frac{p(x,y|c_i)p(c_i)}{p(x,y)} \tag{3}$$

where i is the number of classes. The following rule is applied to choose the class in the example with two classes (Turan and Cengiz, 2022).

$$C(x) = \begin{cases} c_1, p(c_1|x,y) \geq p(c_2|x,y) \\ c_2, p(c_1|x,y) \langle p(c_2|x,y) \end{cases} \tag{4}$$

The following rule can be extended easily for multiclass problems.

2.2.3. Random forest (RF)

RF is one of the most often employed supervised ML algorithms because of its simplicity and versatility, as it can be used to solve classification and regression problems. The RF classifier is part of the ensemble learning method (which provides and constructs several classifiers and aggregates their findings) (Sideris et al., 2019). RF assembles a set of decision trees (DT) to create a forest (Ao et al., 2019). Good predictive outcomes are frequently found when more trees are included in the forest. Individual DT is a set of rules depending on values obtained from the input features tuned to classify all of the training set's examples reliably. If the DT is constructed deeper, it may result in issues such as overfitting due to inconsistencies in the training set. This problem can be resolved using RF, which uses several training samples during the training phase. The number of DTs increases, and the variance decreases, therefore reducing the generalization error and resulting in a robust classifier. Before using RF, the number of trees and each tree's depth level must be tuned. Nonetheless, it is worth noting that as DT increases in depth, the biased power over the training data set increases (Soui et al., 2021). It frequently comes at the expense of a loss in generalization ability. The RF is used to convert the issue into a set of hierarchical requirements, which are expressed as DT. On the other hand, it is not particularly resistant to noisy data. There are three steps in the RF algorithm:

1. We use the original data set to generate ***n*** bootstrap samples (***n*** corresponds to the number of trees built).
2. Growing a classification tree for an individual bootstrap sample by altering it as follows: instead of selecting the best separation between entire predictors, randomly sample ***m*** predictors and choose the best segregation between them in each node.
3. By combining the predictions of the ***n*** trees, additional data predictions can be made (majority vote).

The RF model contains classification trees (C_t), and the final classification model output is as follows:

$$H(x) = \frac{1}{C_t} \sum_{i=1}^{C_t} h_i(x) \tag{5}$$

where $h_i(x)$ are the output of the *i-th* classification tree, and the sample *(x)* output is h_i; therefore, the prediction of RF is the average of all the tree's prediction values.

2.2.4. Extreme gradient boosting (XGBoost)

The XGBoost algorithm is a tree-based ensemble ML model proposed by Chen and Guestrin (Chen and Xgboost, 2016). It is employed in diverse fields and proven by an effective ensemble model. XGBoost works on the "boosting" principle, which combines weak learner prediction with additive training methods to generate a strong learner. (Ullah et al., 2022). Compared to other tree-based ML models (DT and

Table 3
Confusion matrix for evaluating the model's performance.

Actual condition	Predicted condition	
	Positive	Negative
Positive	TP	TN
Negative	FP	FN

RF), XGBoost provides various advantages. First, it employs the second-order Taylor series approximation to improve fitting performance. Second, it includes L2 regularized terms as well as two new strategies to prevent overfitting, namely, shrinkage and column subsampling. Third, it supports parallel computing, which significantly increases the computation speed of boosting algorithms. Fourth, unlike many other existing tree-based ML models, which are either designed for dense data or require special procedures to handle certain circumstances, XGBoost can handle all sparsity patterns in a unified manner, making it ideal, for example, some missing data regarding alternatives. In this study, to obtain higher performance, XGBoost adds L2 regularized terms to prevent overfitting.

$$Z = H(x_i) = \sum_{t=1}^{T} f_t(x_i) \tag{6}$$

where x_i represent the explanatory variables, and $f_t(x_i)$ is each tree output function. To improve the model's predictive capacity, the preferred combination of parameters is needed.

2.3. Shapley additive explanations (SHAP)

Model interpretability is a major problem for ML algorithms that has not received enough attention. Lundberg and Lee proposed the SHAP approach to interpreting the output of the ML model (Lundberg and Lee, 2017). Simple feature importance analysis identifies the critical components; however, it is not apparent how this affects the accuracy of predictions. The most important benefit of SHAP interpretation is that it can reflect the influence of features in every sample and reveal both positive and negative effects (Ullah et al., 0000). Positive features affect the prediction of activity, whereas negative features affect the prediction of inactivity. SHAP is a newly developed interpretation technique that merges optimal credit allocation with local explanation by employing the Shapley value from game theory, where each attribute is referred to as a "contributor". The SHAP value is the model's value assigned to each feature in the sample, which produces a predicted value for each sample (Yang et al., 2021). It can be used to comprehend models in conjunction with various ML models.

In general, the Shapley value in Eq. (7) defines the relevance of a feature i.

$$\phi_i = \frac{1}{|N|!} \sum_{S \subseteq N\{i\}} \frac{|S|!(|N| - |S| - 1)!}{N} [f(S \cup \{i\}) - f(S)] \tag{7}$$

where $f(S)$ corresponds to the XGBoost model's output to be defined by a set S of features, and the whole set of entire features is represented by N. The ultimate contributions $i(\phi_i)$ are calculated as the average contributions over all permutations of a feature set. Finally, the features are independently added to the set, and their significance is shown by the change in the model's output. This method, in particular, considers feature orderings, which influence model output changes when correlated features are present.

2.4. Evaluation metrics

To evaluate and compare the predicted performance of the employed ML algorithms, different classification evaluation metrics were applied. These include accuracy, precision, F1-score, and recall. The confusion matrix for classification problems is made up of four possible scenarios, which are illustrated in Table 3: true positive (TP) count, false positive count (FP), true negative count (TN), and false negatives (FN) count.

Accuracy is defined as the proportion of correctly classified/predicted samples to the entire sample and is presented in eq. (8). Precision measures the total number of correct positive predictions made correctly, whereas recall measures the total number of correct positive predictions that could have been made if all the positive predictions. The formulas for precision and recall were obtained in eqs. (9) and (10). The recall and precision are combined in the F1-score, determined using eq. (11). The F1-score is the harmonic mean of recall and precision. The F1-Score has a range of [0, 1]. It shows how accurate the classifier is (how many instances it correctly classifies) and its robustness (it does not ignore an important number of instances). Our problem is a binary

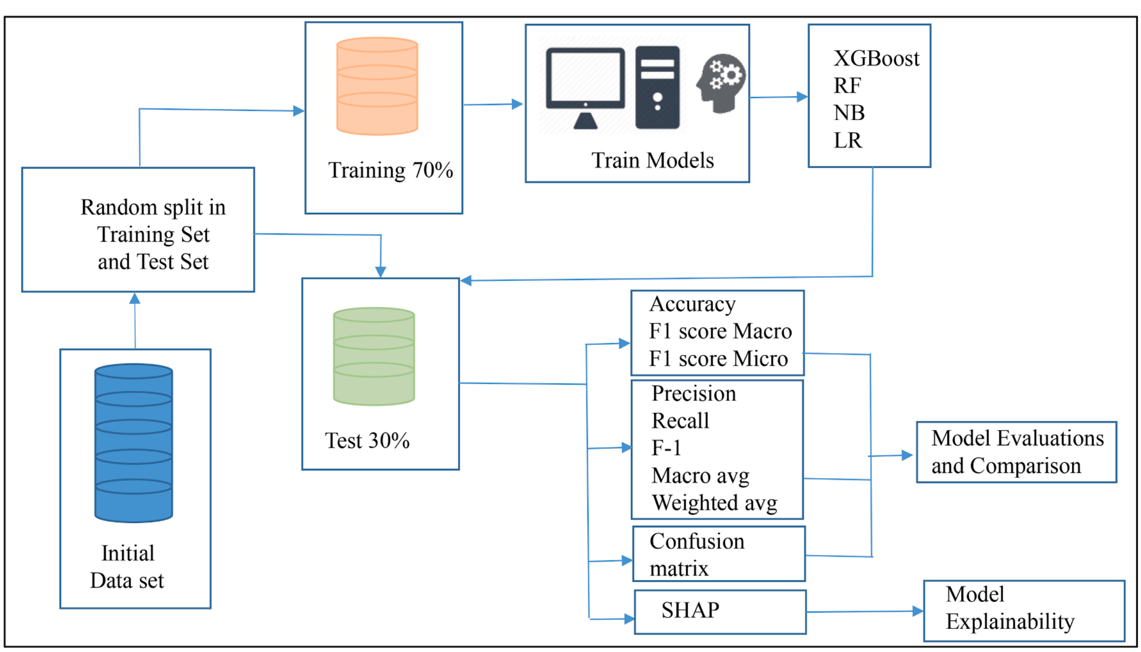


Fig. 1. ML algorithm workflow.

Table 4
Confusion matrix produced for both classes using all ML classifiers.

Models	Actual charging class	Predicted charging class		Correctly classified observations	Incorrectly classified observations
		Normal charging	Fast charging		
XGBoost	Normal charging	27,113	321	27,133	321
	Fast charging	226	2507	2507	226
RF	Normal charging	27,420	24	27,420	24
	Fast charging	2226	507	507	2226
NB	Normal charging	26,039	1395	26,039	1395
	Fast charging	1406	1327	1327	1406
LR	Normal charging	27,156	278	27,156	278
	Fast charging	2629	104	104	2629

Table 5
Summary of all model classification performance.

Models	Accuracy		F1- Macro score		F1-Micro score	
	Train	Test	Train	Test	Train	Test
XGBoost	0.982	0.978	0.945	0.933	0.981	0.977
RF	0.931	0.926	0.823	0.801	0.835	0.823
NB	0.907	0.905	0.717	0.713	0.907	0.905
LR	0.904	0.903	0.751	0.748	0.903	0.902

classification task; we must compute the recall and accuracy rates for each class. To calculate the F1-score, we must take all the different indicator types. The F1-Micro score calculates the total recall and accuracy rate without distinguishing between categories. The F1-macro score calculates each category's recall and accuracy rates and then averages the recall and accuracy rates for each category.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{8}$$

$$Precision = \frac{TP}{TP + FN} \tag{9}$$

$$Recall = \frac{TP}{TP + FN} \tag{10}$$

$$F1 - score = \frac{1}{Precision} + \frac{1}{Recall} \tag{11}$$

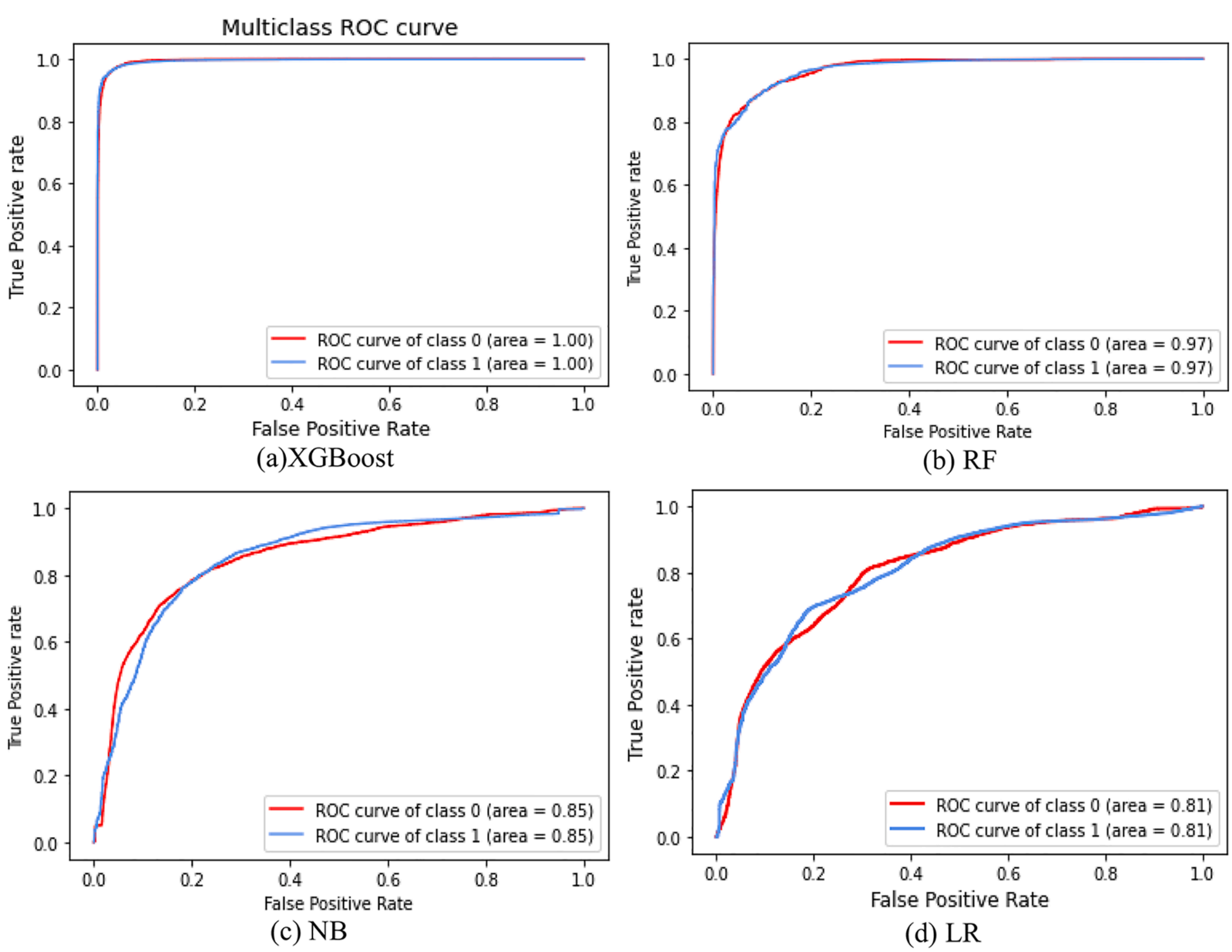


Fig. 2. ROC curve for entire models and different charging classes.

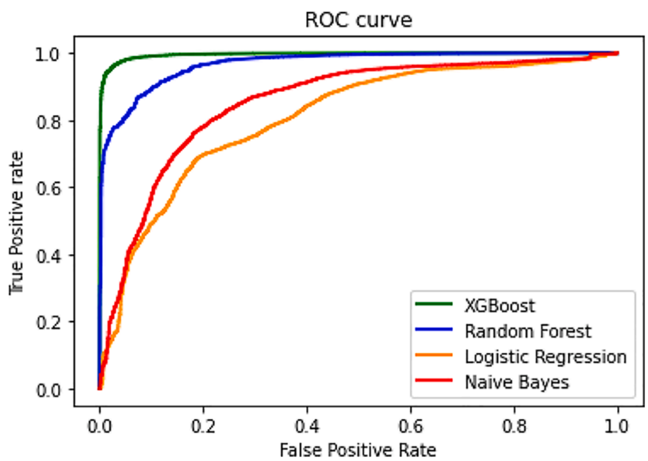


Fig. 3. ROC and AUC metrics for algorithm comparison.

3. Results and analysis

The data set was randomly divided into two parts: training (70 %) and testing (30 %). Four different ML algorithms (LR, NB, RF, and XGBoost) were trained with a training set. Subsequently, ML algorithms were employed to predict the charging station behavior of unknown samples from the test set. The ML algorithm performance was compared using a set of performance metrics. Fig. 1 shows the ML algorithm workflow.

3.1. Model performance evaluation

Table 4 displays the model's prediction performance for each classification algorithm regarding the relevant confusion matrices. In Table 4, confusion matrices exhibit the disparity between actual and predicted observations for both charging classes. The columns in the matrix display the predicted observations for each charging class, and the rows in the matrix indicate actual data for each charging class. The values of the cells across the diagonal reflect accurate predictions, whereas the values of the cells off the diagonal denote inaccurate predictions. The XGBoost model correctly classified 27,113 observations and incorrectly classified 321 observations for the normal charging class. Similarly, for the fast charging class, the numbers of correctly and incorrectly classified observations are 2507 and 226, respectively. Based on the confusion matrix, the XGBoost model more correctly classified observations for both normal and fast charging classes than RF, NB, and LR.

The prediction performance of the proposed classification ML algorithms is summarized in Table 5. Furthermore, in Table 5, the training and test results show no overfitting in the model's result. Accuracy demonstrates the proportion of correctly classified samples of the whole sample. The XGBoost model outperformed the test accuracy (0.978). LR has the lowest accuracy (0.903), followed by NB (0.905) and RF (0.926). Similarly, considering the test F1-macro score (0.933) and F1-micro score (0.977), XGBoost's predictive performance is superior. It is evident from Table 5 that the XGBoost results are more robust than the other models.

Accuracy is the most often used model performance metric, focusing on the effectiveness of prediction outputs. Generally, the model's accuracy is employed as the model's evaluation benchmark. However, the accuracy has some limitations when the distribution of the outcome variable in the sample data is imbalanced. As a result, accuracy is not a perfect performance metric. The receiver operating characteristic (ROC) is introduced as a more accurate metric to measure the effectiveness of a model. To evaluate the predictive performance of the proposed models, the ROC curve and area under the curve (AUC) metrics were used. In Fig. 2, the ROC curves usually have a *Y-axis* for the true positive rate and an *X-axis* for the false positive rate. This suggests that the "ideal" point is in the plot's top left corner, with a false positive rate of zero and an actual positive rate of one. This is not entirely accurate, but it does imply that a greater AUC is usually preferable. The ROC curve "steepness" is also important because it is ideal for increasing the true positive rate while minimizing the false positive rate. The AUC is the area contained by the ROC curve with a maximum value of 1, representing that the result is perfect. A value close to 0.5 suggests that the algorithm generates completely random classification. In Fig. 2, the XGBoost AUC value is 1 for both charging station classes. The RF, NB, and LR AUC values are more than 0.80 for both charging classes. Based on previous studies, with AUC values between 0.9 to 1 and 0.8 to 0.9, the prediction accuracy is excellent and very good, respectively (Tien Bui et al., 2019); an AUC value greater than 0.7 is appropriate for classification problems (Ijaz et al., 2021). Furthermore, the comparison of the ML algorithm by ROC and AUC metrics is displayed in Fig. 3. The XGBoost model performed better than RF, NB, and LR.

While it is essential to obtain an acceptable overall ML model performance, understanding the predictive performance of each charging class is critical to prioritize and select suitable treatment and mitigation measures. Several classification metrics were used to show the model's fitness. Table 6 shows the performance metrics of both charging classes for different models. XGBoost achieved a precision of 0.98 for the normal charging class and 0.89 for the fast charging class. The other

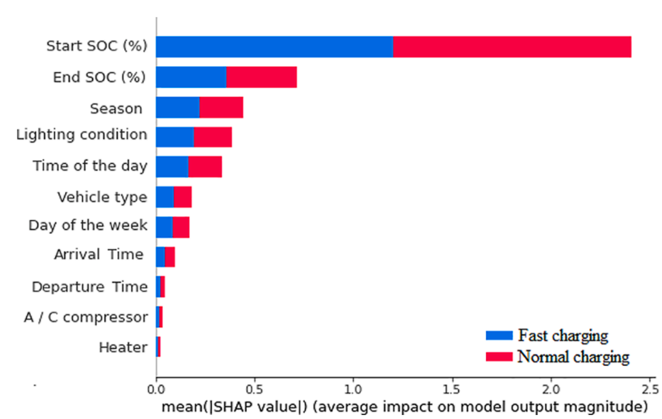
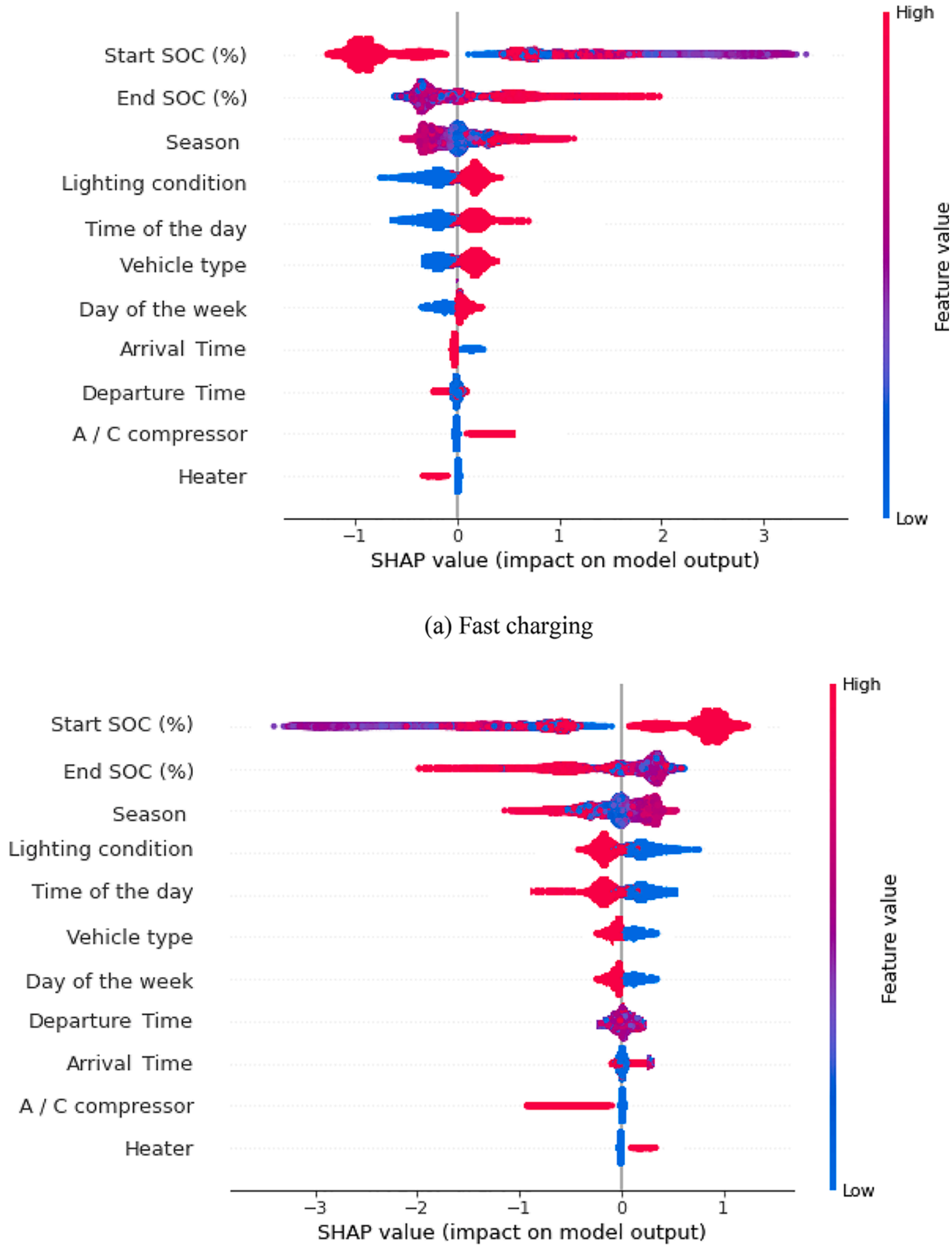


Fig. 4. Summary plot of SHAP values for both classes of the XGBoost model.

Table 6
Performance metrics by both charging station classes.

Models	Charging class	Precision	Recall	F1-score	Macro avg	Weighted avg	Overall accuracy
XGBoost	Normal charging	0.98	0.99	0.99	0.97	0.98	0.98
	Fast charging	0.89	0.92	0.90	0.94	0.98	
RF	Normal charging	0.97	0.99	0.96	0.95	0.59	0.93
	Fast charging	0.72	0.69	0.53	0.93	0.93	
NB	Normal charging	0.95	0.94	0.95	0.91	0.90	0.91
	Fast charging	0.70	0.69	0.70	0.82	0.73	
LR	Normal charging	0.95	0.96	0.95	0.91	0.91	0.90
	Fast charging	0.61	0.55	0.48	0.73	0.70	



(a) Fast charging

(b) Normal charging

Fig. 5. SHAP summary plot.

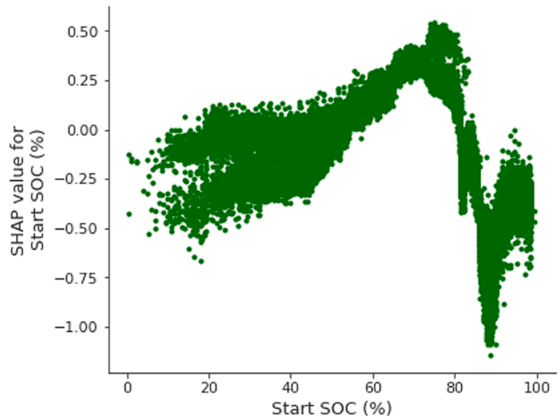
classification metrics had higher performance generated by XGBoost. For the RF, the precision is 0.97 and 0.72 for normal and fast charging classes, respectively. The other classification metrics (NB and LR) achieved acceptable values for both charging classes. Throughout all the performance metrics, the XGBoost model outperformed the other models for both charging classes. Throughout all the performance metrics, the XGBoost model outperformed the other models for both charging classes.

4. Sensitivity analysis

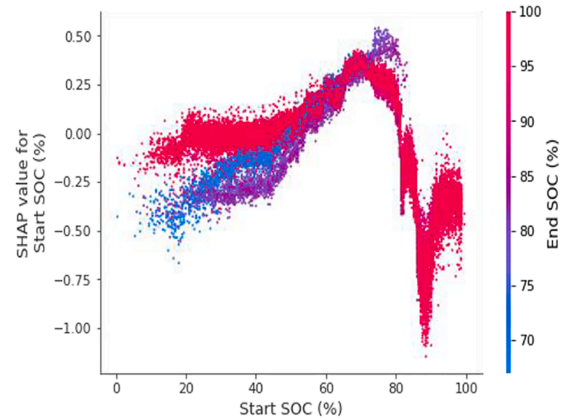
In the previous section, the prediction accuracy of several ML algorithms was examined, and the findings showed that the XGBoost model performed better than the other models. Charging station choice behavior is predicted more accurately if a model can capture the correlation among different features. Therefore, this study further investigates the feature's relative importance in predicting charging station choice behavior. This section shows how the SHAP approach can evaluate the XGBoost results and quantify the impact of certain features.

4.1. Feature importance analysis based on SHAP

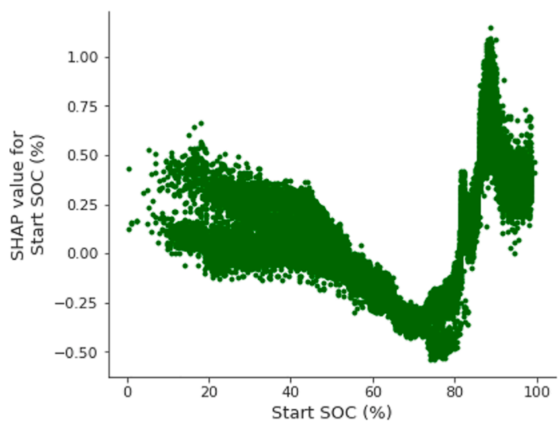
SHAP is used to assess the contribution of independent factors, particularly feature relative importance, to explore the various implications of independent variables on charging classes. A higher value of relative importance suggests that the independent variable has made a more substantial contribution to the estimation of the dependent variable. Fig. 4 shows the average absolute impact of the individual features on the model output magnitude, and the colors blue and red exhibit fast and normal charging classes, respectively. Fig. 4 shows that start SOC, end SOC, season, lighting condition, and time of the day have a high effect on charging classes. The other features show the low effect on



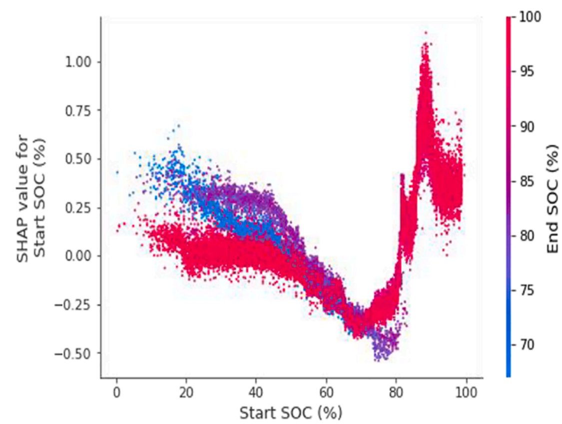
(a) Main effect on Fast charging



(b) Interaction effects on Fast charging



(c) Main effect on normal charging



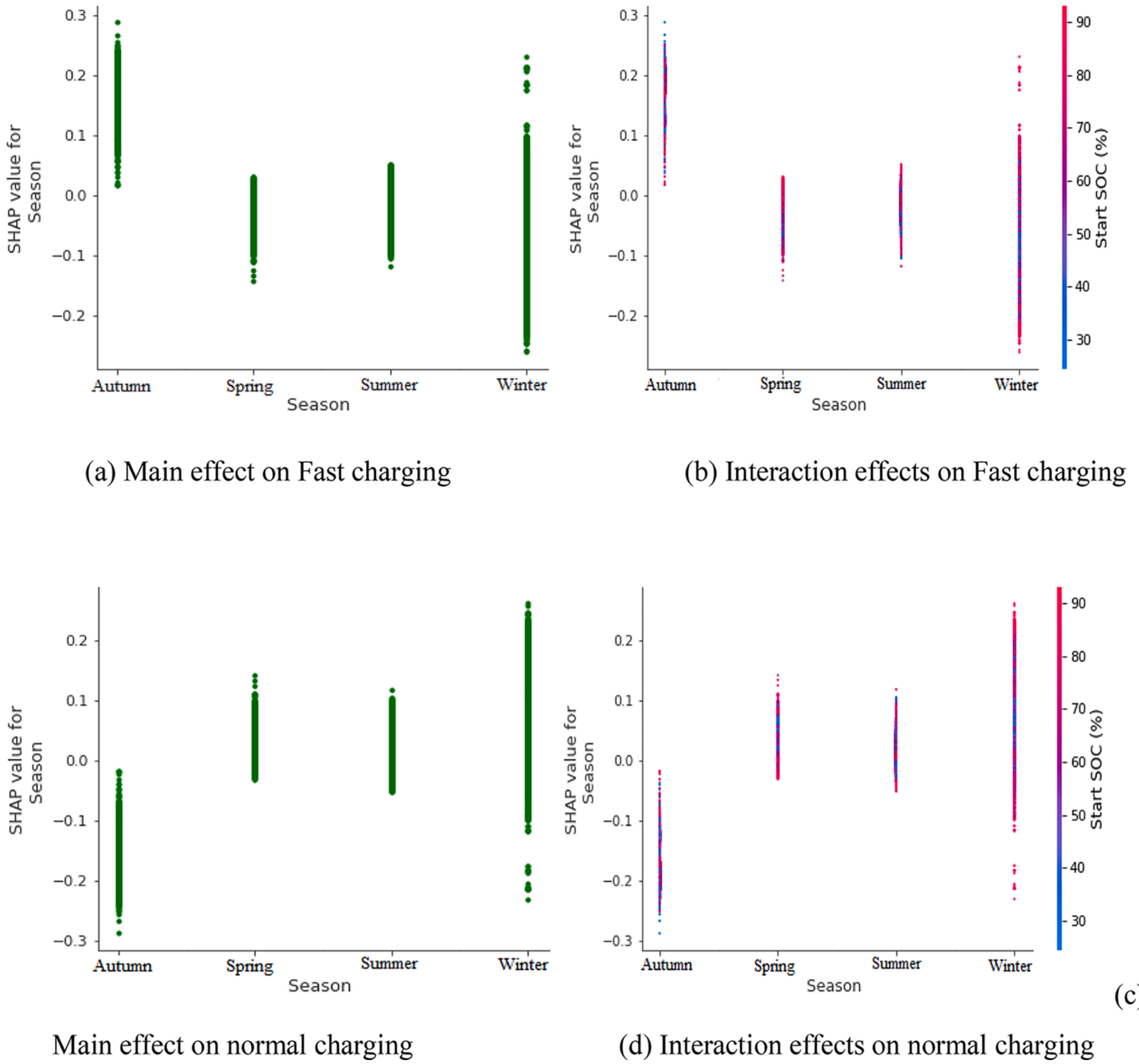
(d) Interaction effects on normal charging

Fig. 6. SHAP main and interaction effects plots of starting SOC.

charging classes.

Fig. 5 displays the SHAP summary plots for both charging classes, which illustrate the overall effect of each feature on the charging class. All the features are ranked in descending order on the primary y-axis, while the SHAP values are shown on the x-axis. The color bar shows how each feature affects the charging class. On the y-axis, the color bar shows the relative importance value of each feature among all the importance of the same feature. Blue shows a lower value, while red indicates a higher value of feature importance. Fig. 5 (a) shows that low start SOC has a positive SHAP, while high start SOC has a negative SHAP value for the fast charging class. Fig. 5 (b) shows that the high start SOC has a positive SHAP value and the low start SOC has a negative SHAP value for the normal charging class. The SHAP value has negative (lower) and positive (higher) effects on each feature of the charging class. The high-end SOC has positive and negative SHAP values for the fast and normal charging classes, respectively. The season (summer and winter) has a positive SHAP value for the fast charging class and a negative SHAP value for the normal charging class.

Under lighting conditions (daytime and nighttime), daytime has positive and negative SHAP values for the fast charging and normal charging classes, respectively, and nighttime has negative SHAP values for fast charging and positive SHAP values for normal charging. The time of the day (peak hours and off peak hours), peak hours has a positive SHAP value for the fast charging class and a negative SHAP value for the normal charging class. Vehicle type (commercial vehicle and private vehicle), commercial vehicles have positive and negative SHAP values for fast and normal charging, respectively, while private vehicles have negative SHAP values for fast charging and positive SHAP values for normal charging. The day of the week (weekdays and weekend), weekdays have a positive SHAP value for fast charging and a negative SHAP value for normal charging; likewise, the weekend has negative and positive SHAP values for fast and normal charging, respectively. The SHAP result revealed that a start SOC percentage of less than 30 % may be a turning point for the fast charging choice rather than the normal charging choice. Conversely, the start SOC percentage of more than 80 % is more likely to choose normal charging than fast charging. If the percentage of end SOC is more than 70 %, this could be the turning point for fast charging choice, while those with an end SOC percentage less than 40 % are more likely to prefer normal charging choice. This indicates that the charging choice behaviors vary based on the SOC percentage. Charging during summer and winter preferred the fast charging choice rather than the normal charging choice. Furthermore, the SHAP result shows that charging in the daytime is more inclined to choose the fast charging choice, and nighttime charging prefers the normal charging choice. Interestingly, charging in peak hours is more likely to result in a fast charging choice, and charging in off-peak hours is more likely to result in a normal charging choice. Similarly, charging on weekdays and weekends preferred fast charging and normal charging, respectively. Furthermore, commercial vehicles prefer fast charging, and private vehicles prefer normal charging. Fig. 5 (a-b) shows that the



(a) Main effect on Fast charging (b) Interaction effects on Fast charging

(c) Main effect on normal charging (d) Interaction effects on normal charging

Fig. 7. SHAP main and interaction effects plots of season.

high-effect features for both charging classes are start SOC, end SOC, season, lighting condition, and time of the day. The rest of the features show a low effect on both charging classes.

The explanatory variables used for the SHAP analysis impact charging behavior (as shown in Fig. 5) and are expected to provide valuable insights for practical policy implications. It has been reported that charging EVs is a function of demand-based behavior (Xu et al., 2017). For example, considering SOC indicators, it is expected that a lower initial SOC is linked with a high possibility of fast charging behavior. In contrast, a high initial SOC allows more freedom for EV drivers to use home-based normal chargers rather than to be forced to use public charging stations. This is because EVs with large-capacity batteries in cars that are fully charged are less prone to driving range anxiety. Variable lighting conditions and time of the day have also been reported to influence EV charging behavior (Ullah et al., 2022). Studies have shown that the probability of home charging is relatively high after midnight until 7 am, while drivers tend to avoid rush hours during the day (Zoepf et al., 2013). Similarly, the usage of (A/C and heater) varies during the day and nighttime travel; thereby, the charging consumed and charging behavior choice are impacted accordingly. Vehicle type is another significant predictor influencing charging behavior. Commercial vehicles with large batteries and high charging demand are expected to use public charging stations more often than private vehicles (Sun et al., 2016). Likewise, considering the day of the week indicator, the typical travel pattern of EVs is different during weekdays and weekends, and hence, the charging behavior is supposed to be different. The variables vehicle departure and arrival time also play a significant role in the selection and preference for charging stations (fast or normal) at a given time. For example, arrival and departure rates at specific times discourage users from charging their vehicles at public and normal charging stations. Finally, excessive A/C compressors and heater usage are expected to drain the EV battery quickly, so drivers usually have to keep a close eye on the remaining driving range and accordingly plan for the next charging event. The number of days between the start of the current charging event and the start of the next trip is also an important factor in deciding whether to choose a normal or fast charging station. Users are expected to perform normal charging if time permits; otherwise, fast charging will be their preferred choice.

4.2. Main and interaction effects

The SHAP dependence plots incorporate the main and interaction effects of several essential features, with the interaction effects depicted as the vertical dispersion of SHAP values in different colors. This dispersion depicts how the two variables influence the target variable simultaneously. We showed each main charging class and the interaction effects of the two features in parallel. In Figs. 6-10, the main and interaction effects of different features for both charging classes are

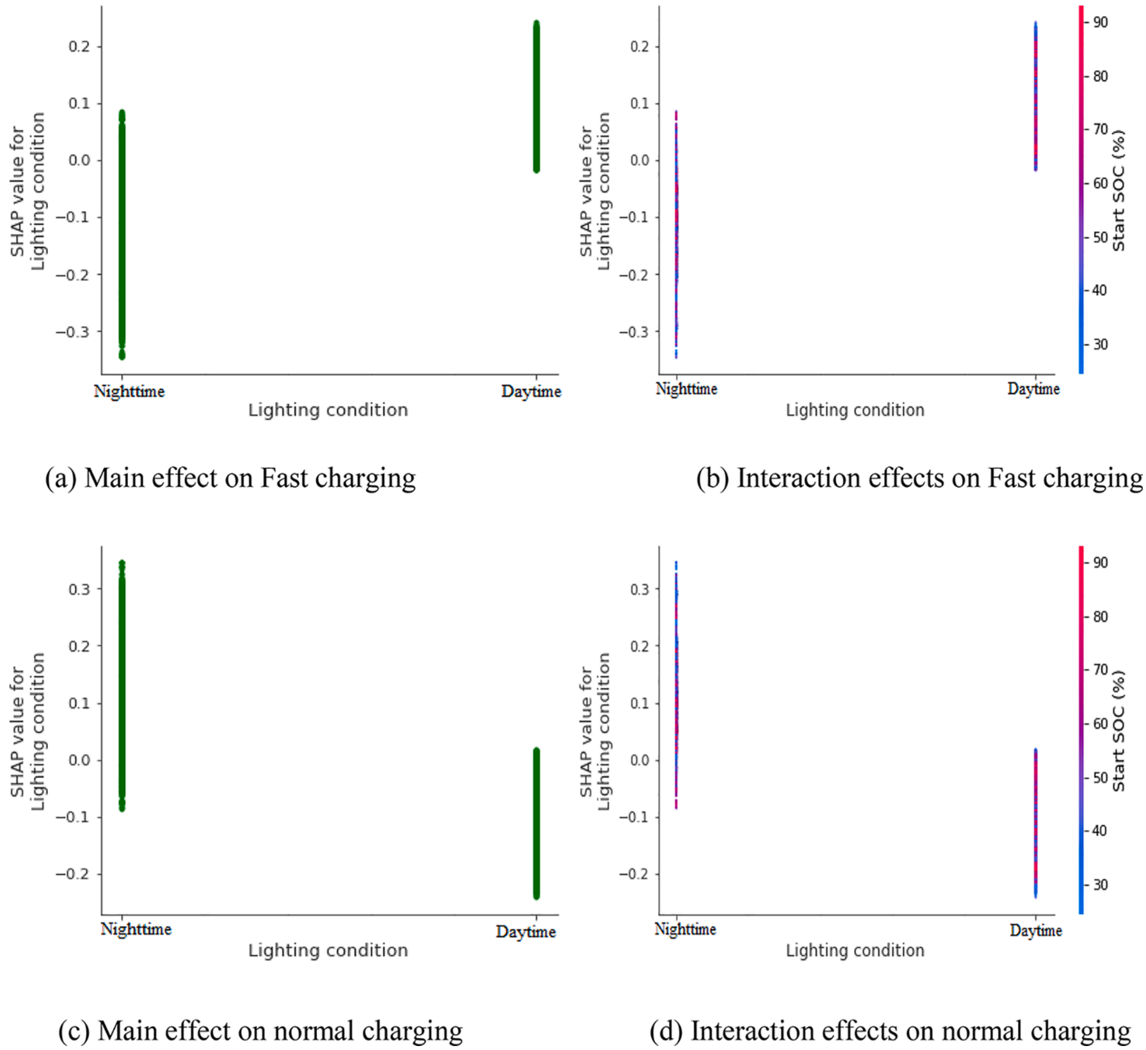
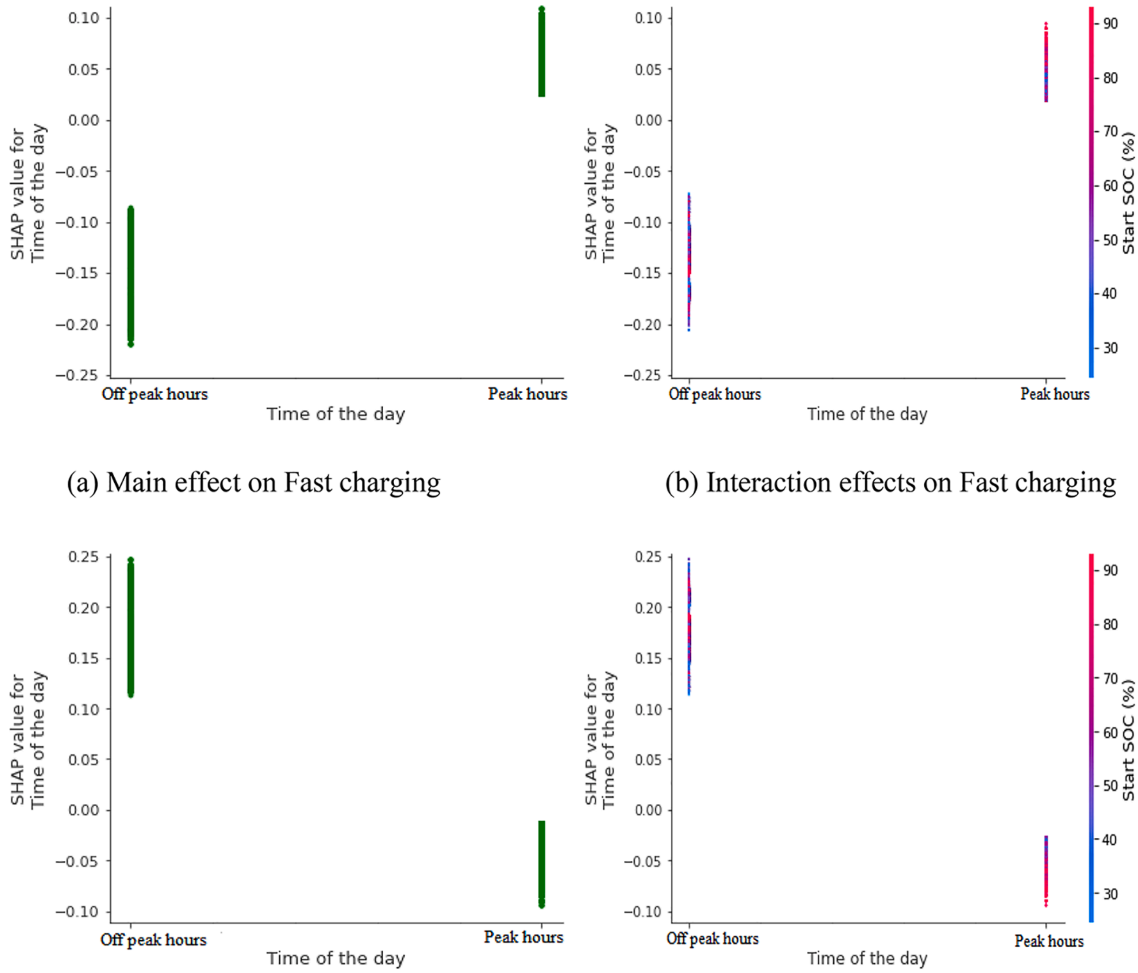


Fig. 8. SHAP main and interaction effects plots of lighting conditions.

plotted separately side by side in this section. The left Figs. 6-10 (a) and (c) show the main effects, while the right Figs. 6-10 (b) and (d) show the interaction effects, and the colors red and blue indicate the low and high start and end SOC, respectively. In Fig. 6, the main effects of the start SOC and the interaction effects of the start SOC and end SOC are extracted and compared for both charging classes. In Fig. 6 (a), the start SOC increases from 15 % to 80 %, and the SHAP value increases and stabilizes at a positive value (i.e., higher impact on charging choice behavior). This indicates the preferred fast charging choice if the SOC percentage is 15 to 80 %. Increasing the start SOC from 80 % shows a negative SHAP value (i.e., lower impact on charging choice behavior) for fast charging choice. If the SOC percentage increases from 80 %, it suggests that normal charging is the preferable choice. On the other hand, the vertical dispersion in Fig. 6 (b) suggests an increase in the start SOC corresponding to the high end SOC. The opposite conclusion holds for Fig. 6 (c) and (d) for normal charging. This finding shows that if the SOC percentage is less than 80 %, fast charging is preferable compared to normal charging.

The main effects of different seasons and the interaction effects of season and start SOC are shown in Fig. 7 (a-d). Autumn has positive and negative SHAP values for normal and fast charging, respectively, and spring, summer, and winter have positive SHAP values for both charging. The interaction effect of seasons shows the low start SOC for autumn and high start SOC for spring, summer, and winter for both charging choices. These results found that charging in autumn is a preferable fast charging choice if the start SOC is between (30–50 %) compared to normal charging and the spring, summer, and winter have the same preferences for both charging choices. The main effects of lighting conditions and the interaction effect of lighting conditions and start SOC are shown in Fig. 8 (a-d). Fig. 8 (a) displays the nighttime charging (negative SHAP value) and daytime charging (positive SHAP value) for fast charging choice, where the vertical dispersion in Fig. 8 (b) implies that a low and high start SOC leads to daytime and nighttime charging for fast charging choice. The same phenomenon is a contradiction for the normal charging choice in Fig. 8 (c) and (d). The findings show that daytime charging is more inclined to choose fast charging, and nighttime charging tends to be normal charging. Fig. 9 (a-d) shows the main effects of the day and the interaction effect of the time of the day and start SOC. Charging in peak and off peak hours has a positive and negative SHAP value for fast charging, respectively. Conversely, for normal charging, the peak hours and off peak hours charging have a negative and positive SHAP value, respectively. The vertical dispersion in Fig. 9 (b) and (d) implies that low and high start SOC leads to charging in peak hours/off peak hours for both charging. These outcomes indicate that charging in peak hours is more inclined to fast charging and off peak hours to tend to normal charging. The same conclusion holds for Fig. 10.



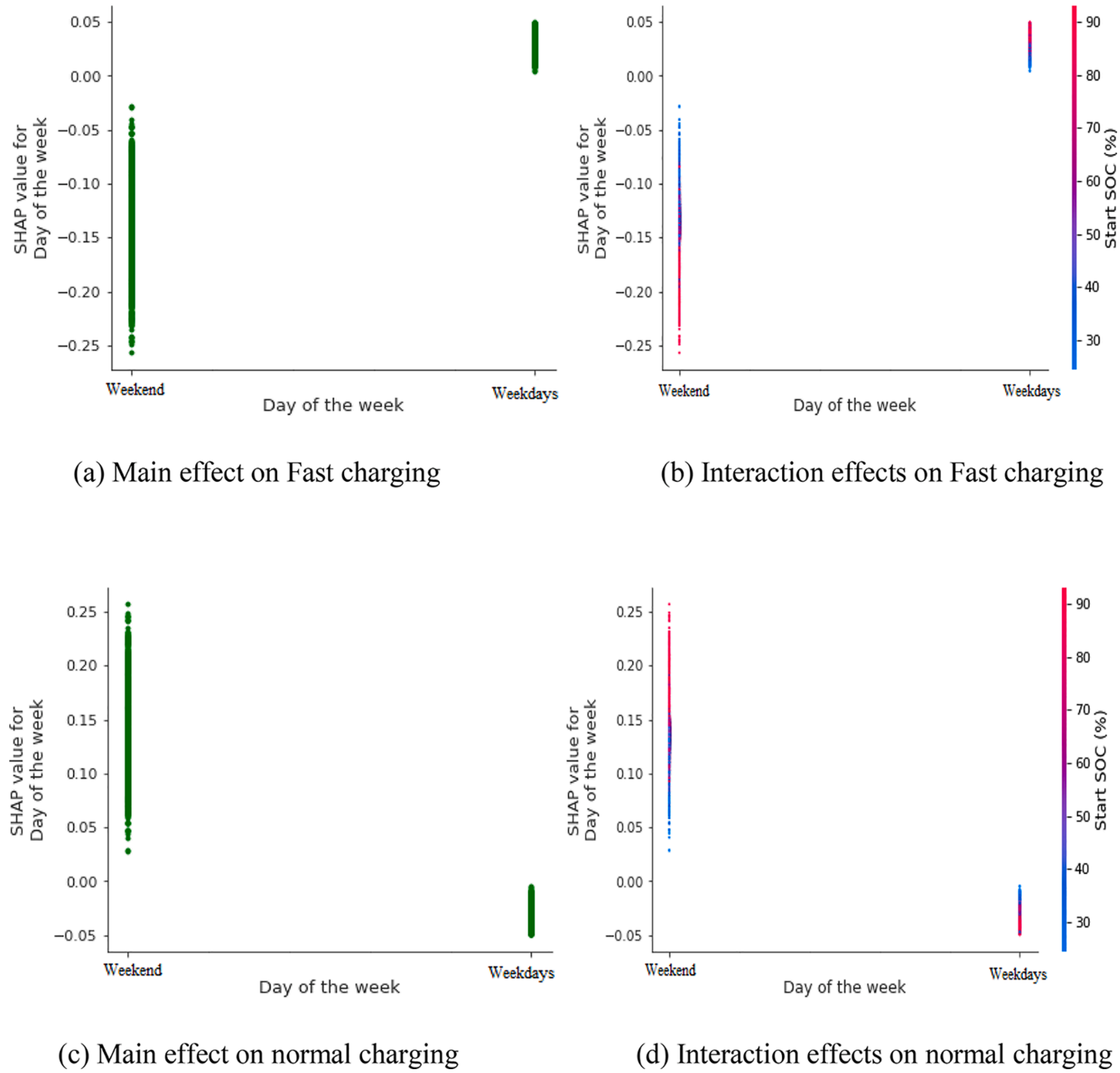
(a) Main effect on Fast charging (b) Interaction effects on Fast charging

(c) Main effect on normal charging (d) Interaction effects on normal charging

Fig. 9. SHAP main and interaction effects plots of time of the day.

5. Significance and comparison of the employed ML model's interpretation

Developing a suitable ML application to determine the appropriate algorithm among a considerable number of potential algorithms is challenging. However, researchers are often more at ease and confident using a model they comprehend and interpret effectively. Furthermore, each method has its own hyperparameters and a feature outside the model whose value cannot be estimated from the data (Bas et al., 2021). There are various methods to evaluate the interpretability of particular ML techniques. LR is the common probabilistic-based statistical model used to solve the classification problem. To estimate probabilities, LR commonly uses a logistic function, often known as the mathematically defined sigmoid function. As opposed to linear regression, the assigned weights in LR do not affect the probability of the class outcome linearly. The major drawback of LR is the assumptions of linearity between variables. Similar to LR, in the NB model, the numeric weight is adjusted initially to several input variables, which are progressively changed. The performance of the NB model may be affected by its strong assumptions regarding feature independence. However, these models criticized the implication behind the black box. Tree-based ML approaches such as RF and XGBoost are sequential models that coherently combine a number of simple tests, where every test evaluates a nominal attribute against a set of probable outputs or a numeric feature against a threshold value (Kashifi et al., 2022). The logical rules given by tree-based models are frequently easier to comprehend than the connection weights generated by other models. The XGBoost models have good interpretability, fast training, support parallel learning, and high efficiency in data handling, but the RF algorithm has a complex model structure, low learning speed, and sensitivity to variations in the data (Kashifi et al., 2022). XGBoost minimizes a regularized L1 and L2 objective function comprising a convex loss function and a model complexity penalty term. Moreover, regularization through both L1 and L2 avoids overfitting. To make the final prediction, training continues iteratively by adding new trees that predict the residuals or errors of prior trees, which are then integrated with prior trees (Chen and Xgboost, 2016). In addition, XGBoost is a nonlinear boosting model that performs well in most data sets, either binary classification or multiclassification and imbalanced problems. In statistical models, the significance of the input variables is based on the t value for a specific confidence level. In these models, the influence of features is determined by negative and positive signs. In contrast, ML models rank input variables based on feature scores, information gain, connection weights, and SHAP values. ML model interpretability is critical for understanding the logic behind ML model prediction or decision making. Performance measurements such as prediction accuracy may not be sufficient to describe practicality. Traditional ML models are inherently interpretable since they have a simple structure (Molnar, 2020). However, due to the complexity of ensemble models and deep neural networks, it is difficult to understand the behavior of the models (Du et al., 2019). These models require a post hoc approach to improve



(a) Main effect on Fast charging

(b) Interaction effects on Fast charging

(c) Main effect on normal charging

(d) Interaction effects on normal charging

Fig. 10. SHAP main and interaction effects plots of day of the week.

explainability by extracting data from the models' input and output (Molnar, 2020). There are frequent tradeoffs between the accuracy of a prediction and its interpretability. More interpretable models may produce less accurate predictions; especially compared to less interpretable models (Du et al., 2019). Thus, although statistical models have been employed to classify charging choice behavior, their prediction accuracy has been demonstrated to be inferior to complex ML models. Lundberg developed SHAP to enhance the interpretability of the ML model (Lundberg and Lee, 2017). SHAP is a model-agnostic ML interpretability method that may be used for any ML model. It also provides a local explanation for each data sample as well as a global explanation for the entire data set. It could be challenging to comprehend ML's behavioral predictions for charging choice; nonetheless, the ML model with the SHAP approach in this area appears to be rapidly developing at the convergence of statistical analysis to investigate patterns in complex data sets. Furthermore, in the current big data-driven era, the importance of theoretical investigation and implementations has resulted in significant breakthroughs in analytical methods, with ML becoming highly significant.

6. Conclusion

EVs have been an important research topic in recent decades because they can help reduce global air pollution and are a promising choice for contributing to greenhouse gas emission reduction goals. This study aimed to develop an ML model with a SHAP approach to predict EV charging station choice behavior as a function of several predictor variables. In this study, a first-time ML model with a SHAP approach was utilized for EV charging station choice behavior prediction. Unlike previous shallow structure approaches, it is unclear how the features affect prediction outcomes; the ML models with a SHAP approach were proposed to interpret the ML model's output. The data utilized in this research were procured from real-world normal and fast charging event data of 500 EVs over two years from Japan. ML algorithms (LR, NB, RF, and XGBoost) were used to train and test the data set. The obtained results demonstrated that the predictive accuracy of all the proposed ML algorithms on the training set is better than that on the test set for both charging classes. The final prediction made on the test data set. The confusion matrices were used to check the predictive performance of the models. The confusion matrices show that the XGBoost model outperformed other ML models regarding charging class prediction. The ROC curve and AUC metrics were used to evaluate the predictive performance of the individual charging classes. The XGBoost obtained AUC value is 1 for both charging classes. The RF, NB, and LR AUC values are more than 0.80 for both charging classes. This shows that the ML algorithm achieved good performance for both charging classes.

Nevertheless, XGBoost had superior prediction performance.

Many ML algorithms, including neural networks, have been criticized for their inability to provide actual prediction dependencies on features. Even though some tree-based algorithms prioritize features, we can only find a linear relationship between predictions and features. The SHAP approach establishes an interpreter for all features to interpret and visualize the result. Further SHAP dependence plots illustrate a considerably more complicated marginal dependence of each feature on the projected outcome of an ML model. The most significant features for both charging are the start SOC, end SOC, season, lighting condition, and time of the day. Furthermore, the SHAP result shows that the individual features affect charging choice behavior. The SHAP result revealed that the low start SOC preferred the fast charging choice; increasing the start SOC is more likely to result in normal charging. Increasing the percentage of end SOC favors choosing fast charging, while the low end SOC is more inclined to choosing normal charging. In regard to charging in the summer and winter, the fast charging choice is preferred, while charging in autumn and spring favors the normal charging choice. Furthermore, charging during the day is more likely to choose fast charging, although charging at night prefers normal charging. Charging during peak hours is more likely to favor fast charging, whereas charging during off-peak hours favors normal charging. Based on these findings, employing the SHAP approach, the XGBoost model has been proven to effectively capture the impact of influencing factors on the charging choice. The proposed study's findings provide valuable guidelines and recommendations to practitioners and decision-makers in deploying adequate EV charging infrastructure in metropolitan areas.

There are a few potential limitations of this study that should be acknowledged. The data set used in this study is relatively old. Furthermore, the data lack detailed drivers' social characteristics, battery type, charging price, and traffic conditions, which are expected to significantly impact charging behavior. The EV performance and charging parameters evolve with time. In the future, it will be necessary to use recent data sets (containing a rich set of explanatory variables), which may provide further insight into users' charging behavior and optimal deployment of charging infrastructure.

CRediT authorship contribution statement

Irfan Ullah: Conceptualization, Methodology, Software, Visualization, Investigation, Data curation, Formal analysis, Writing - original draft, Writing - review & editing. **Kai Liu:** Conceptualization, Methodology, Investigation, Funding acquisition, Writing - review & editing. **Toshiyuki Yamamoto:** Data curation, Writing - review & editing. **Muhammad Zahid:** Methodology, Software, Visualization. **Arshad Jamal:** Investigation, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The authors thank the Japan Automobile Research Institute (JARI) for providing the dataset. The corresponding author would also like to acknowledge support funding from National Natural Science Foundation of China (Grant No. 51378091 and 71871043).

References

Adornato, B., Patil, R., Filipi, Z., Baraket, Z., Gordon, T., 2009. Characterizing naturalistic driving patterns for Plug-in Hybrid Electric Vehicle analysis. In: IEEE Vehicle Power and Propulsion Conference, pp. 655–660. https://doi.org/10.1109/VPPC.2009.5289786.

Ai, S., Chakravorty, A., Rong, C., 2018. Household EV charging demand prediction using machine and ensemble learning. In: 2018 IEEE International Conference on Energy Internet (ICEI). IEEE, pp. 163–168.

Ao, Y., Li, H., Zhu, L., Ali, S., Yang, Z., 2019. The linear random forest algorithm and its advantages in machine learning assisted logging regression modeling. J. Petroleum Sci. Eng. 174, 776–789.

Azadfar, E., Sreeram, V., Harries, D., 2015. The investigation of the major factors influencing plug-in electric vehicle driving patterns and charging behaviour. Renew. Sustain. Energy Rev. 42, 1065–1076. https://doi.org/10.1016/j.rser.2014.10.058.

Bas, J., Zou, Z., Cirillo, C., 2021. An interpretable machine learning approach to understanding the impacts of attitudinal and ridesourcing factors on electric vehicle adoption. Transportation Lett. 1–12.

Berdichevsky G, Kelty K, Straubel JB, Toomre E. The Tesla Roadster Battery System Tesla Motors. Recuperado de: Http://Large Stanford Edu/Publications/Coal/References/Docs/Tesla Pdf 2006.

Chakraborty, D., Hardman, S., Tal, G., 2022. Integrating plug-in electric vehicles (PEVs) into household fleets-factors influencing miles traveled by PEV owners in California. Travel Behav. Soc. 26, 67–83.

Chen T, Guestrin C. Xgboost: A scalable tree boosting system. Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining, 2016, p. 785–94.

Chung, Y.-W., Khaki, B., Li, T., Chu, C., Gadh, R., 2019. Ensemble machine learning-based algorithm for electric vehicle user behavior prediction. Appl. Energy 254, 113732.

De Sousa L, Silvestre B, Bouchez B. A combined multiphase electric drive and fast battery charger for electric vehicles. 2010 IEEE Vehicle Power and Propulsion Conference, IEEE; 2010, p. 1–6.

Du, M., Liu, N., Hu, X., 2019. Techniques for interpretable machine learning. Commun. ACM 63 (1), 68–77.

ElMenshawy, M., Massoud, A., 2020. Modular Isolated DC-DC Converters for Ultra-Fast EV Chargers: A Generalized Modeling and Control Approach. Energies 13, 2540.

ElMenshawy, M., Massoud, A., 2020. Hybrid Multimodule DC-DC Converters for Ultrafast Electric Vehicle Chargers. Energies 13, 4949.

Franke, T., Krems, J.F., 2013. Understanding charging behaviour of electric vehicle users. Transport. Res. F: Traffic Psychol. Behav. 21, 75–89. https://doi.org/10.1016/j.trf.2013.09.002.

Frendo, O., Graf, J., Gaertner, N., Stuckenschmidt, H., 2020. Data-driven smart charging for heterogeneous electric vehicle fleets. Energy and AI 1, 100007.

Harrington, P., 2012. Machine learning in action. Simon and Schuster.

Ijaz, M., lan, L., Zahid, M., Jamal, A., 2021. A comparative study of machine learning classifiers for injury severity prediction of crashes involving three-wheeled motorized rickshaw. Accid. Anal. Prev. 154, 106094.

Kashifi, M.T., Jamal, A., Kashefi, M.S., Almoshaogeh, M., Rahman, S.M., 2022. Predicting the travel mode choice with interpretable machine learning techniques: A comparative study. Travel Behav. Soc. 29, 279–296.

Liao, F., Molin, E., van Wee, B., 2017. Consumer preferences for electric vehicles: a literature review. Transp. Rev. 37 (3), 252–275.

Liu, K., Gao, H., Liang, Z., Zhao, M., Li, C., 2021. Optimal charging strategy for large-scale electric buses considering resource constraints. Transp. Res. Part D: Transp. Environ. 99, 103009.

Liu, K., Gao, H., Wang, Y., Feng, T., Li, C., 2022. Robust charging strategies for electric bus fleets under energy consumption uncertainty. Transp. Res. Part D: Transp. Environ. 104, 103215.

Lu, Y., Li, Y., Xie, D.a., Wei, E., Bao, X., Chen, H., Zhong, X., 2018. The application of improved random forest algorithm on the prediction of electric vehicle charging load. Energies 11 (11), 3207.

Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. Proceedings of the 31st international conference on neural information processing systems, 2017, p. 4768–77.

Majidpour M, Qiu C, Chu P, Gadh R, Pota HR. A novel forecasting algorithm for electric vehicle charging stations. 2014 International Conference on Connected Vehicles and Expo (ICCVE), IEEE; 2014, p. 1035–40.

Molnar, C., 2020. Interpretable machine learning. Lulu. com.

Pan, L., Yao, E., MacKenzie, D., 2019. Modeling EV charging choice considering risk attitudes and attribute non-attendance. Transportation Research Part C: Emerging Technologies 102, 60–72.

Rahman, I., Vasant, P.M., Singh, B.S.M., Abdullah-Al-Wadud, M., Adnan, N., 2016. Review of recent trends in optimization techniques for plug-in hybrid, and electric vehicle charging infrastructures. Renew. Sustain. Energy Rev. 58, 1039–1047.

Rawson, M., Kateley, S., 1999. Electric vehicle charging equipment design and health and safety codes. SAE Trans. 3256–3262.

San Román, T.G., Momber, I., Abbad, M.R., Sánchez Miralles, Á., 2011. Regulatory framework and business models for charging plug-in electric vehicles: Infrastructure, agents, and commercial relationships. Energy Policy 39 (10), 6360–6375.

Shibl, M., Ismail, L., Massoud, A., 2021. Electric Vehicles Charging Management Using Machine Learning Considering Fast Charging and Vehicle-to-Grid Operation. Energies 14, 6199.

Sideris, N., Bardis, G., Voulodimos, A., Miaoulis, G., Ghazanfarpour, D., 2019. Using random forests on real-world city data for urban planning in a visual semantic decision support system. Sensors 19, 2266.

Soui, M., Mansouri, N., Alhamad, R., Kessentini, M., Ghedira, K., 2021. NSGA-II as feature selection technique and AdaBoost classifier for COVID-19 prediction using patient's symptoms. Nonlinear Dyn. 106 (2), 1453–1475.

Sun, X.-H., Yamamoto, T., Morikawa, T., 2016. Fast-charging station choice behavior among battery electric vehicle users. Transp. Res. Part D: Transp. Environ. 46, 26–39.

Thomas, C.E., 2009. Fuel cell and battery electric vehicles compared. Int. J. Hydrogen Energy 34 (15), 6005–6020.
Tien Bui, D., Shirzadi, A., Chapi, K., Shahabi, H., Pradhan, B., Pham, B., Singh, V., Chen, W., Khosravi, K., Bin Ahmad, B., Lee, S., 2019. A hybrid computational intelligence approach to groundwater spring potential mapping. Water 11 (10), 2013.
Turan, S.C., Cengiz, M.A., 2022. Ensemble learning algorithms. J. Sci. Arts 22 (2), 459–470.
Ullah I, Liu K, Yamamoto T, Zahid M, Jamal A. Prediction of electric vehicle charging duration time using ensemble machine learning algorithm and Shapley additive explanations. International Journal of Energy Research n.d.
Ullah, I., Liu, K., Yamamoto, T., Zahid, M., Jamal, A., 2021. Electric vehicle energy consumption prediction using stacked generalization: an ensemble learning approach. Int. J. Green Energy 18 (9), 896–909.
Ullah, I., Liu, K., Yamamoto, T., Al Mamlook, R.E., Jamal, A., 2022. A comparative performance of machine learning algorithm to predict electric vehicles energy consumption: A path towards sustainability. Energy Environ. 33 (8), 1583–1612.
Ullah, I., Liu, K., Yamamoto, T., Shafiullah, M., Jamal, A., 2022. Grey wolf optimizer-based machine learning algorithm to predict electric vehicle charging duration time. Transportation Letters 1–18.
van Heuveln, K., Ghotge, R., Annema, J.A., van Bergen, E., van Wee, B., Pesch, U., 2021. Factors influencing consumer acceptance of vehicle-to-grid by electric vehicle drivers in the Netherlands. Travel Behav. Soc. 24, 34–45.
Walsh, J., Heazlewood, I.T., Climstein, M., 2019. Regularized linear and gradient boosted ensemble methods to predict athletes' gender based on a survey of masters athletes. Model Assisted Stat. Appl. 14 (1), 47–64.
Wang, C., Yang, J., Liu, N., Mao, Y., 2011. Study on siting and sizing of battery-switch station. In: 2011 4th International Conference on Electric Utility Deregulation and Restructuring and Power Technologies (DRPT). IEEE, pp. 657–662.
Weng, J., Yang, D., 2015. Investigation of shipping accident injury severity and mortality. Accid. Anal. Prev. 76, 92–101.
Westin, K., Jansson, J., Nordlund, A., 2018. The importance of socio-demographic characteristics, geographic setting, and attitudes for adoption of electric vehicles in Sweden. Travel Behav. Soc. 13, 118–127.
Xiong, Y., Chu, C., Gadh, R., Wang, B., 2017. Distributed optimal vehicle grid integration strategy with user behavior prediction. In: IEEE Power & Energy Society General Meeting. IEEE, pp. 1–5.
Xu, Z., 2017. Forecasting electric vehicle arrival & departure time on UCSD campus using support vector machines. University of California, San Diego.
Xu, M., Meng, Q., Liu, K., Yamamoto, T., 2017. Joint charging mode and location choice model for battery electric vehicle users. Transp. Res. B Methodol. 103, 68–86.
Yang, C., Chen, M., Yuan, Q., 2021. The application of XGBoost and SHAP to examining the factors in freight truck-related crashes: An exploratory analysis. Accid. Anal. Prev. 158, 106153.
Yang, Y., Yao, E., Yang, Z., Zhang, R., 2016. Modeling the charging and route choice behavior of BEV drivers. Transp. Res. Part C: Emerg. Technol. 65, 190–204.
Yang, Y., Tan, Z., Ren, Y., 2020. Research on factors that influence the fast charging behavior of private battery electric vehicles. Sustainability 12, 3439.
Youn, E., Jeong, M.K., 2009. Class dependent feature scaling method using naive Bayes classifier for text datamining. Pattern Recogn. Lett. 30 (5), 477–485.
Zhao, X., Yan, X., Yu, A., Van Hentenryck, P., 2020. Prediction and behavioral analysis of travel mode choice: A comparison of machine learning and logit models. Travel Behav. Soc. 20, 22–35.
Zhu, R., Hu, X., Hou, J., Li, X., 2021. Application of machine learning techniques for predicting the consequences of construction accidents in China. Process Saf. Environ. Prot. 145, 293–302.
Zoepf, S., MacKenzie, D., Keith, D., Chernicoff, W., 2013. Charging choices and fuel displacement in a large-scale demonstration of plug-in hybrid electric vehicles. Transp. Res. Rec. 2385 (1), 1–10.