


## Article

# Forecasting Charging Point Occupancy Using Supervised Learning Algorithms

Adrian Ostermann <sup>1,2,\*</sup> , Yann Fabel <sup>1</sup>, Kim Ouan <sup>3</sup> and Hyein Koo <sup>1,3</sup><sup>1</sup> FfE Munich, 80995 München, Germany; yfabel@ffe.de (Y.F.); hkoo@ffe.de (H.K.)<sup>2</sup> School of Engineering and Design, Technical University of Munich (TUM), 80333 München, Germany<sup>3</sup> Faculty of Informatics, Technical University of Munich (TUM), 80333 München, Germany; kim.ouan@tum.de

\* Correspondence: aostermann@ffe.de

**Abstract:** The prediction of charging point occupancy enables electric vehicle users to better plan their charging processes and thus promotes the acceptance of electromobility. The study uses Adaptive Charging Network data to investigate a public and a workplace site for predicting individual charging station occupancy as well as overall site occupancy. Predicting individual charging point occupancy is formulated as a classification problem, while predicting total occupancy is formulated as a regression problem. The effects of different feature sets on the predictions are investigated, as well as whether a model trained on data of all charging points per site performs better than one trained on the data of a specific charging point. Reviewed studies so far, however, have failed to compare these two approaches to benchmarks, to use more than one algorithm, or to consider more than one site. Therefore, the following supervised machine-learning algorithms were applied for both tasks: linear and logistic regression, k-nearest neighbor, random forest, and XGBoost. Further, the model results are compared to three different naïve approaches which provide a robust benchmark, and the two training approaches were applied to two different sites. By adding features, the prediction quality can be increased considerably, which resulted in some models performing better than the naïve approaches. In general, models trained on data of all charging points of a site perform slightly better on median than models trained on individual charging points. In certain cases, however, individually trained models achieve the best results, while charging points with very low relative charging point occupancy can benefit from a model that has been trained on all data.

**Keywords:** electric vehicles; charging points; occupancy; supervised learning; forecasting

**Citation:** Ostermann, A.; Fabel, Y.; Ouan, K.; Koo, H. Forecasting Charging Point Occupancy Using Supervised Learning Algorithms. *Energies* **2022**, *15*, 3409. <https://doi.org/10.3390/en15093409>

Academic Editor: Caiping Zhang

Received: 28 February 2022

Accepted: 4 May 2022

Published: 6 May 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The European Union (EU) aims to become climate neutral by 2050 to fulfill its commitment to the Paris Agreement, keeping global warming well below 2 °C and making efforts to limit it to 1.5 °C [1]. Even though the emissions of the European electricity sector were 39% lower in 2019 compared to their 1990 levels, even completely converting the generation of electricity to renewable sources will not be sufficient to achieve the goal of net-zero greenhouse gas (GHG) emissions [2]. To achieve climate neutrality, both a strong increase in energy efficiency and far-reaching electrification in the energy-consuming sectors of industry, buildings (heating and cooling), and transport need to be realized, in addition to exclusively climate-neutral electricity generation [3–5]. As opposed to the other sectors, GHG emissions from the transport sector increased between 2013 and 2019 [6]. In 2018, more than 12% of EU GHG emissions were caused by passenger cars [6]. All previous emission reductions (achieved via, e.g., more efficient engines, or new fuels such as E10) have been negated by an increased volume of traffic and growing numbers of vehicles with comparatively high fuel consumption, which makes climate neutrality in transport seem particularly challenging to achieve. One of the most promising means of reducing emissions in the transport sector is the adoption of electric vehicles (EV), provided that the

electricity consumed is generated from renewable energy sources [7]. Market penetration of EVs, however, is faced with the ‘chicken-and-egg’ problem: while the absence of charging infrastructure limits the proliferation of EVs, investments in infrastructure require higher charging demand from EVs to become attractive [8]. While the EU has neither passed a law requiring certain numbers or market shares of EVs, nor committed to non-binding targets, the European Commission’s current revised proposal for the Fit for 55 package strengthens the current 2030 average fleet-wide emissions from newly registered vehicles from  $-37,5\%$  to  $-55\%$  relative to a 2021 benchmark. In addition, the proposal calls for a complete phase-out of internal combustion engine vehicles by 2035, with a 100% reduction in greenhouse gas emissions from new cars and vans sold from 2035 onwards [9]. Meanwhile at the national level, the German government has set a target of ten million registered EVs and one million publicly accessible charging points by 2030 as part of the Climate Protection Program 2030 (German: Klimaschutzprogramm 2030) [10]. This goal aligns with the European Alternative Fuels Infrastructure Directive (AFID), which calls for a 10:1 ratio of EVs to publicly accessible charging points [11].

However, further obstacles also must be addressed to ensure the ramp-up of electromobility and thus to ultimately achieve the climate targets in the transport sector. Several studies on EV adoption [12–15] highlight that one of the main barriers to purchasing an EV is insufficient or unavailable charging infrastructure. Further, due to the sparsely developed infrastructure, EV users must plan their charging processes more deliberately than users of conventional vehicles. This circumstance is accompanied by another frequently cited barrier to EV adoption [16–18]: range anxiety, i.e., the fear of running out of power before reaching an available charging station. Webpages and apps such as those in [19–21] help EV users plan their route by either displaying accessible charging points or taking the vehicle range and possible necessary charging stops into account. The webpage Go-TO-U [22] even shows current occupancy status, although this is only helpful for users actively looking for a free charging station at the given moment. Only 1 out of 9 charging points in the EU is fast-charging ( $>22$  kW). Consequently, with normal charging points ( $\leq 22$  kW) being dominant in the EU, charging typically takes several hours [23]. In order for EV users to plan their next charging session, forecasting charging point occupancy is essential. Predicting charging point occupancy is also advantageous for scheduling charging point maintenance tasks during times of low usage.

EV users heavily rely on the precise prediction of the battery range. Therefore, proper battery models predicting battery states, e.g., state-of-health remaining useful life, state-of-charge or state-of-energy are crucial. Many existing studies in the literature examined different battery modeling methods [24–27]. The models include physics-based electrochemical models, analytical models, stochastic models, and electrical-circuit based models. However, according to Tomasov et al. electrochemical-based models as well as equivalent circuit ones are best suited for usage in transport applications [24]. Tremblay’s battery model is specifically designed for EV application and can accurately represent the dynamic behavior of the battery [28]. Tremblay et al. obtained battery model parameters by evaluating just three points from the manufacturer’s discharge curve in steady state. The study by Zhang et al. was conducted on optimization of Tremblay’s battery model parameters for plug-in hybrid EV applications [29]. Since the original method by Tremblay et al. is error prone, Zhang et al. proposed to use a novel quantum-behaved particle swarm optimization (QPSO) parameter-estimation technique to estimate the model parameters [28,29]. Another model employed by [30–32] is the Volterra model, which can be used to approximate complex nonlinear dynamics inherent to the battery. Hu et al. proposed to integrate a linear double-capacitor model, an equivalent circuit model, with a data-based Volterra model to build a physics-informed data-driven model for lithium-ion batteries, called Volterra double-capacitor model [30]. They demonstrated this new model’s high accuracy via experimental validation. Sidorov et al. presented both an example for battery modeling for renewable energy applications and an adaptive approach to solve load leveling problem with storage [32]. In both cases, they used a dynamic evolutionary

model based on the first kind Volterra integral equation. They came to the conclusion that the systems of Volterra equations can model energy storage systems combining the different energy storage technologies.

In addition to the research on modeling energy storage systems in the aforementioned studies, there are a number of studies addressing optimized planning of the charging infrastructure and selection of the appropriate location. Pagany et al. [33] provide a broad range of charging infrastructure planning methods and give an extensive review of various methods regarding the required input data, theoretical approaches, and maturity. Metais et al. [34] reviewed different model options for planning EV charging infrastructure, which can be divided into two categories depending on the objective: minimizing the cost of charging infrastructure for a given level of service, or maximizing the service provided for a given cost. Additionally, node, path, and tour-based approaches can be distinguished. For instance, Micari et al. [35] proposed a methodology to calculate the required number of charging stations for EVs and find their optimal location in a road network. The methodology consists of a two-level model using the demand (i.e., the flow of EVs) and the supply (i.e., the road network where the charging stations will be positioned). It was validated on a test road network and subsequently applied to the Italian highway network. Guo et al. [36] used the fuzzy technique for order preference by similarity to ideal solution (TOPSIS) method to select the optimal EV charging station site and built an evaluation index system as a selection criterion.

Other studies focus on the existing charging infrastructure and their usage. Viswanathan et al. [37] analyzed the charging infrastructure in San Diego and developed an assessment model to calculate the demand of public charging points. One finding of the study is that even though San Diego had enough chargers for the charging demand of existing EV's, the public charging distribution network was neither well-designed nor effectively used. Flammini et al. [38] used a dataset from the Dutch research center ElaadNL containing 2900 public charging points with more than 400,000 charging transactions from the year 2015. They analyzed the charging behavior on weekdays, Saturday, and Sunday and derived probability density functions describing the plug-in and -out events. One finding is that 50% of the charging times were less than 4 h, while the idle time of a charging point depends on the geographic location of the charging point and lasts on average for 4 h. By collecting empirical data from almost 27,000 charging points in Germany, Hecht et al. [39] investigated the usage intensity of the charging infrastructure: With 15% to 25%, the overall usage intensity of the stations in Germany is relatively low but is strongly dependent on the weekday and time of day. They further demonstrated that a high share of charging events last 8 to 10 h, indicating that EVs are parked significantly longer at charging stations than the actual charging duration.

Another widely addressed field is charging demand and flexibility. Almaghrebi et al. [40] aimed to predict session charging demand at public charging stations using supervised machine-learning regression algorithms on a dataset from Nebraska, USA. They used four different models (linear regression, gradient boosting, random forest, and support vector machine), which only provided moderate accuracy, accounting for roughly 50% of the variance in user behavior. They showed that for this dataset, the primary statistic of predictive value is the user's average demand for past sessions. However, they stated that using the same framework for a smaller area or a single charging station, where input features have a higher correlation to the energy demand, could result in better predictions. Xiong et al. [41] modelled EV user behavior based on EV user charging records collected in public charging facilities by combining statistical analysis and machine-learning approaches to predict day-ahead EV load on the UCLA campus. This was carried out by applying unsupervised clustering in combination with a multilayer perceptron to historical charging records. Their results show that their model achieves good performance for charging control scheduling and EV load forecasting. Lucas et al. [42] performed supervised machine-learning regression on a public charging station dataset from The Netherlands to estimate idle time of an EV after charging. They compared the performance of gradient

boosting, random forests, and XGBoost using most of the available features in the dataset, and showed that idle time can be estimated with a fair degree of accuracy for the studied dataset. Sørensen et al. [43] used field data from Norway from more than 6800 charging sessions to describe charging behavior and to estimate charging and flexibility potential for residential EVs. The results were validated using data from installed advanced metering systems. Considerably good results in residential EV charging flexibility estimation were achieved when private parking spaces were equipped with their own charging point. Gerritsma et al. [44] and Sadeghianpourhamami et al. [45] analyzed the flexibility of EV charging demand using charging sessions from The Netherlands. Gerritsma et al. [44] found that 59% of the aggregated EV charging demand can be deferred by more than 8 h and 16% by more than 24 h. Sadeghianpourhamami et al. [45] grouped the charging sessions into three behavioral clusters based on arrival and departure time (1) “park to charge”, (2) “charge near home” and (3) “charge near work”, and quantified the flexibility for a load-flattening and load-balancing scenario. They further found that the flexibility of all clusters was influenced by seasonal changes and weekends and that the sessions in the “charging near home” cluster were better suited to shifting charging demand into nighttime.

In the field of predicting charging point occupancy, the following studies have been published [37–40]. Bikcora et al. [46] used a dataset from ElaadNL representing public charging stations and performed a general linear model, namely logistic regression (LogR), to forecast the availability of each charging point. The researchers found that simple model structures, possibly consisting of hourly indicator variables, day-lagged and week-lagged measured variables, were shown to be sufficient for the disaggregated level of charging points. However, the study only focused on public charging stations and does not describe data preparation, perform any parameter optimization, or apply and compare other linear or non-linear models. Tian et al. [47] combined real-time GPS data from electric taxis in Shenzhen with historic charging events for each taxi to develop a real-time charging-station-recommendation system. First, they predicted the current operational state of each taxi and recharge intention. When an electric taxi driver sends a recharge request to the system, the driver is recommended a charging station where the total cost of charging is most likely to be minimal. One disadvantage of this method is that the charging behavior and the current GPS data of the vehicle are necessary. Furthermore, the charging intentions of other drivers must be known to the system in order to calculate the waiting time. Majidpour et al. [48] used data from the UCLA campus and performed four time-series models: historical average (naïve approach), k-nearest neighbor (KNN), weighted KNN and lazy learning. They further proposed a time-weighted dot-product-based nearest neighbor to improve their results. Like the work presented herein, Motz et al. [49] used the Adaptive Charging Network (ACN) data and investigated the influence of different feature sets and whether it is better to use data of the individual charging points or of all charging points. The ACN project provides access to real electric vehicle charging sessions from different sites. Motz et al., however, only considered the workplace site from the can data. Further, they only used a linear model (LogR) and omitted a comparison to a benchmark model.

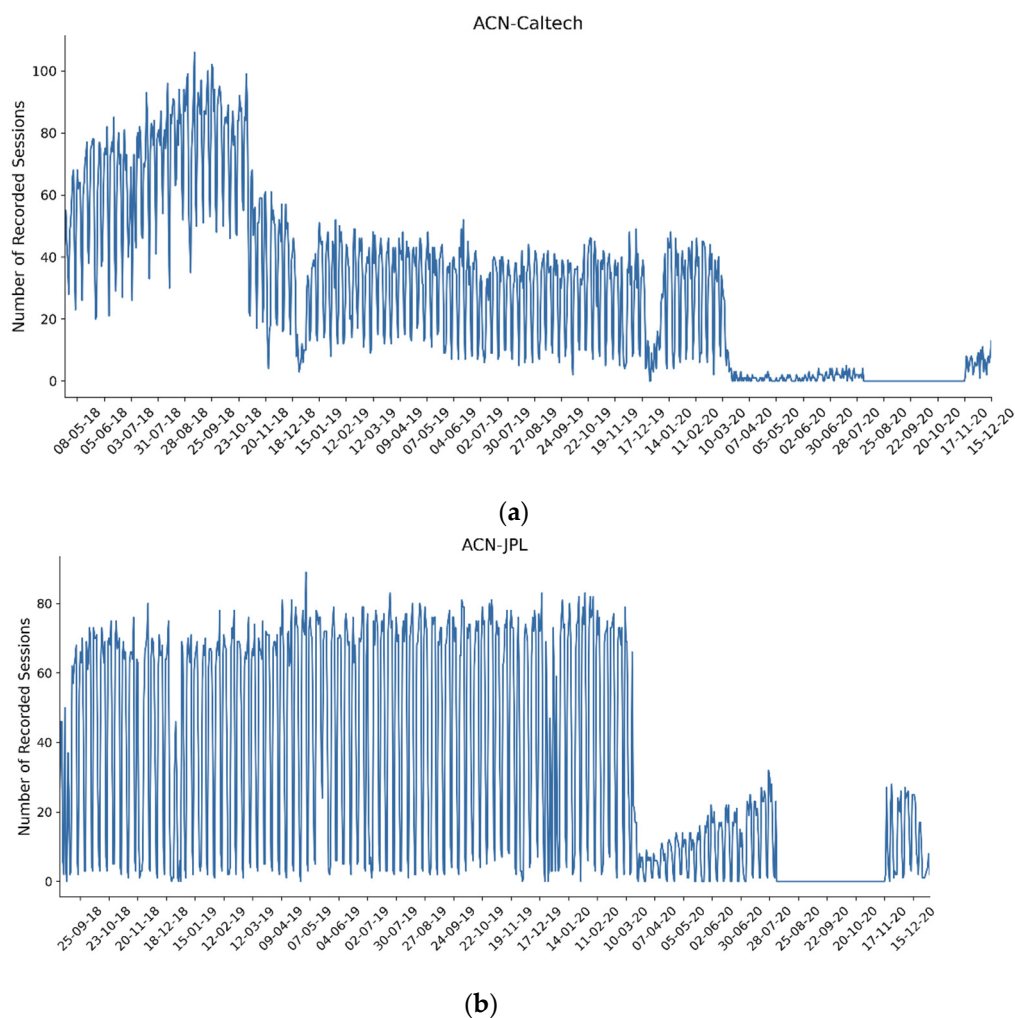
Our contribution to this topic is that we compare workplace and public charging sites and apply different classification models (linear and non-linear) to predict charging point occupancy, and compare them to three benchmark models. Therefore, we first describe the used data and the data preparation in detail. After this, we perform hyperparameter optimization to find the optimal model configuration. This is carried out for models trained on individual charging points and models trained on all charging point data for each site. We compare two different feature sets and their influence on the prediction, as well as the different training modes. Finally, we implement four different regression models (linear and non-linear) to predict the total number of occupied charging points at each site.

## 2. Materials

For our analysis, we use the publicly available Adaptive Charging Network (ACN) dataset [50]. This dataset consists of two different sites: Caltech and JPL. The Caltech

University site has 54 charging points, or electric vehicle supply equipment (EVSE), in one campus garage, which is located next to the university gym. The charging points are open to the public and are often used by non-Caltech drivers. We consider Caltech a public charging site. The JPL site is a national research lab with 50 EVSEs and is only open to employees with a key card, and therefore is characterized as a workplace charging location. Lee et al. [50] describe the dataset and how it is collected, while Lee et al. [51] and Lee et al. [52] give more details on the charging facility and adaptive algorithm. In addition to charging station data, the dataset also contains user data (scheduled departure time and requested energy), which users can enter via an app. In total, the dataset contains 15 features, including the “userInputs” which contains a list of eight more features. Like Motz et al. [49], we use the same three features: station ID, connection time, and disconnect time, because most of the excluded features are either redundant or do not have any information for our purpose. These three features are the minimum requirements for charging event datasets. Limiting our analyses to these three features therefore allows the presented method to be applied and the results compared to other datasets.

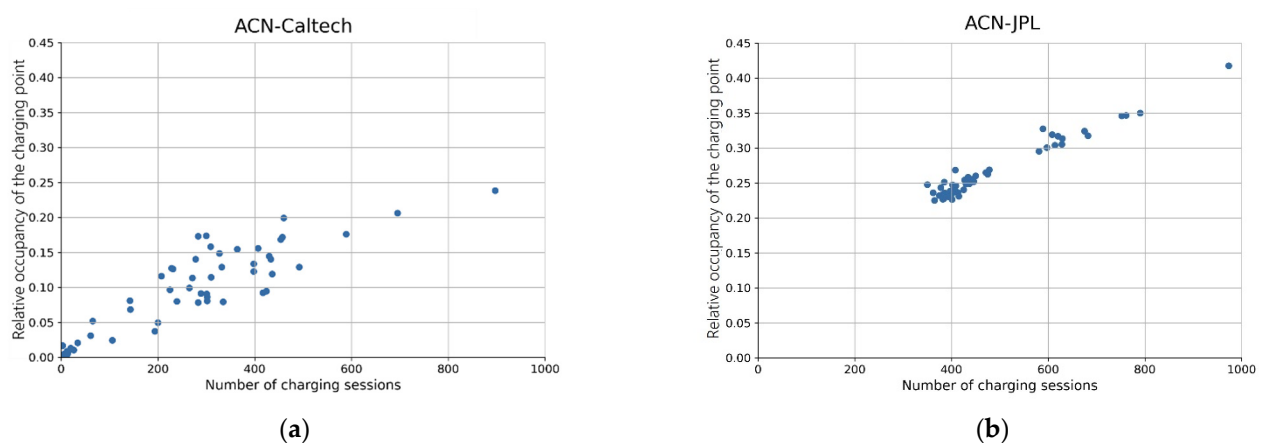
Figure 1 shows the distribution of the number of recorded sessions per day from the beginning of available data at Caltech (top) and JPL (bottom). Both Caltech and JPL have more than 25,000 registered charging sessions. The start date of the available data for the Caltech site is 4 April 2018, while for JPL it is 9 September 2018. The data were collected on 12 December 2020, representing the end date.



**Figure 1.** (a) Number of recorded sessions for Caltech; (b) Number of recorded sessions for JPL.

As seen in Figure 1, Caltech (top) shows three and JPL (bottom) two different usage periods. As stated by Lee et al. [50], parking was completely free at the beginning at Caltech and, due to an issue with site configuration, for approximately half of the EVSEs at JPL until 1 November 2018. This change is reflected in the significant reduction in the recorded sessions per day at the Caltech site, while at JPL there is no pattern change, likely due to the demand for charging overshadowing any price sensitivity, according to Lee et al. [50]. After 1 November 2018, the recorded number of sessions at Caltech is around half the amount compared to JPL. The winter recess lasted from 15 December 2018 to 6 January 2019, and from 14 December 2019 to 5 January 2020. Both periods are observable in the data due to drops in recorded sessions. Caltech shows a stronger decrease in recorded sessions during these periods. This suggests that the site is mostly dominated by students, who leave the campus for holidays, while JPL is still used by the employees. However, we still consider Caltech a public site since no key card is needed nor is any other restriction to use the charging points present. The next notable pattern change at both sites is at the beginning of March 2020, marking the beginning of the COVID-19 pandemic. In order to consider a comparable period in each case, we use the data for JPL from 5 September 2018 to 29 February 2020, and from 1 December 2018 to 29 February 2020 for Caltech, excluding the COVID-19 period at both sites. After the paid parking switch, it took about one month for the charging behavior to normalize, which is why we exclude this period as well.

Figure 2 shows the relative occupancy for each charging point per number of charging sessions for Caltech (a) and JPL (b). The relative occupancy is calculated for each charging point by the total occupation time divided by the time period. Therefore, the relative occupancy rate provides information on how frequently a charging point is used during the examined period. Some stations at Caltech (a) have a relative occupancy below 0.05 with less than 200 charging sessions, while with increasing number of charging sessions, the relative occupancy rises. To put this into perspective, over the considered period of 456 days or 10,944 h, these charging points were only used for a total of around 23 days, or 547 h. Based on 200 charging sessions, this results in an average session length of 2.7 h. Caltech shows a more heterogeneous distribution across the EVSEs compared to JPL (b) and has an overall occupancy rate of 0.14. Almost all EVSEs at JPL are as frequently used as the highest at Caltech, with an overall relative occupancy of 0.27. All EVSEs at JPL have more than 350 charging sessions and mean relative occupancy of 0.27.



**Figure 2.** Relative occupancy for each charging point (a) at Caltech; (b) at JPL.

### 3. Methods

In this chapter, we first describe the methodology for the two predictive models we developed. One performs regression analysis to predict the amount of occupied charging stations at each site and time interval. To predict the occupancy status of individual charging stations at the two sites (Caltech, JPL) for a specified time interval, the second model uses a classification method. The classification algorithm further employs two different

approaches. The first approach was designed to predict the occupancy of individual charging points. A separate model for each individual charging point is trained and tested on point-specific data. For the second approach, the models are trained and tested with the data from all charging points of the corresponding site, leading to a single predictive model for each site. In the following, we refer to the first approach as mode ‘individual’ and to the second one as mode ‘all’. In this way, we aim to answer the question of whether it is more beneficial to use all available data of one site or not.

### 3.1. Data Preparation

To begin with, we present the process of data preparation, give a short overview of the utilized classification and regression models, discuss the hyperparameter tuning, and complete this chapter with an outline of the utilized metrics in this work for evaluating and comparing the results.

As described in Section 2, data were excluded from the original datasets due to external events that had a significant impact on the distribution of charging sessions. Since these events are not represented within the input data, negative consequences regarding the learning effect can be assumed, which is why the affected time periods were removed entirely. In the end, the dataset of the Caltech site consists of 456 days, while the JPL dataset contains 543 days of charging data. The original dataset holds timestamps of the connection and disconnection time for respective EVSEs. However, to predict the occupancy status of individual charging points, the exact timestamp of a future charging session is irrelevant (and not feasible to predict). Rather, the focus is on the time frame in which a charging session is expected. Therefore, we first grouped the data by EVSE-ids and defined time intervals of 15 min, hence converting the data into time series. For each time interval, we then defined the dependent (target) variable. In case of the classification, the dependent variable is a Boolean value ‘occupancy’, which is set to true if a charging session occurs within the interval. For the regression task, the dependent variable refers to the number of occupied EVSE within the respective site.

As mentioned in Section 2, we did not use additional features that the ACN dataset holds due to reasons of comparability. Rather, we split and extended the time interval feature, i.e., the timestamp, into multiple separate features: ‘month’, ‘day of month’, ‘weekday’, ‘holiday’, ‘bridging day’ (a single workday between a public holiday and weekend), and ‘interval’. The features ‘holiday’ and ‘bridging day’ were determined based on the public holidays of the state of California (where the respective charging stations are located), whereas ‘interval’ describes the number of quarter hours since the beginning of the day, hence an integer value between 0 and 95.

Moreover, we analyzed the impact of including two additional features describing the occupancy status of a respective EVSE for a specific interval on the previous day and the previous week (24 h before, 7 days before). We refer to this feature set as ‘augmented’. For mode ‘individual’, we created one instance of time series per site and for each charging point ordered over time. For mode ‘all’, we combined all individual charging point time series into one per site ordered over time. Therefore, using the mode ‘all’ dataset, the models were trained on one single dataset containing the time series of all charging points. In the next step, the JPL and Caltech datasets were split into a training, validation, and test set. Thus, overfitting can be detected, and the generalization capabilities of the predictive models can be analyzed. Since the datasets consist of time series data, we split the data into temporally contiguous parts with the validation and the test sets comprising the two recent months of data.

In the final step of the data preparation, we used cyclic feature encoding to account for periodic patterns in time-based features. The problems with cyclical data for machine-learning algorithms are the jump discontinuities. Mahajan et al. found that LogR and Linear regression (LinR) benefit from using cyclic feature encoding and suffer when ordinal encoding was used [53]. Further, they found that classification and regression trees suffer

from the choice of one-hot encoding, and might be more robust towards raw cyclical features. Distance-based models such as KNN also profit from this encoding method.

A simple approach defined by [54] involves dividing a feature into a sine and cosine part, for a feature  $x \in [0, x_{max}]$  (e.g., weekday  $\in [0, 6]$ ), this results in:

$$x_{\sin} = \sin\left(2 \times \pi \times \frac{x}{x_{max}}\right) \quad (1)$$

$$x_{\cos} = \cos\left(2 \times \pi \times \frac{x}{x_{max}}\right) \quad (2)$$

After applying these transformations to all corresponding features, the standard feature set for classification and regression looks as shown in Table 1.

**Table 1.** Feature description for standard feature set.

Feature Name	Value Range	Cyclic Feature Encoding
Month	[0, 11]	yes
Day of month	[0, 31]	yes
Weekday	[0, 6]	yes
Interval	[0, 95]	yes
Holiday	{0, 1}	no
Bridging day	{0, 1}	no

The additional features for the augmented feature set are listed in Table 2.

**Table 2.** Feature description for additional features in augmented feature set.

Use-Case	Feature Name	Value Range	Cyclic Feature Encoding
Classification	Previous week	{0, 1}	no
	Previous day	{0, 1}	no
Regression	Previous week	[0, max(num_evse)]	no
	Previous day	[0, max(num_evse)]	no

### 3.2. Supervised Learning Algorithms

Supervised learning algorithms learn from a training set of labeled examples, where for each observation of the predictor measurement(s)  $x_i$ ,  $i = 1, \dots, n$  there is an associated response measurement  $y_i$ , [55]. The algorithm or model developed then aims to map the response to the predictors by optimizing a given objective function to make accurate predictions [55]. After training, the quality of the prediction is evaluated to subsequently apply the learned patterns to unknown data with the same characteristics as the training set. Two different types of supervised learning can be distinguished depending on the type of response: classification and regression. Classification refers to predicting a qualitative response or categorical classes of new instances based on previous observations. The labels of these classes are unique values that can be understood as the group membership of the instances. According to the number of classes, one differentiates between binary classification when the response measurement takes only two values or multiclass classification for more than two. If the predictor(s) is continuous or quantitative, regression models are used [55]. The following sections describe the models underlying this research.

#### 3.2.1. Naïve Model

For benchmarking our machine-learning-based models, we evaluate three naïve models. For two of them, we use a target value from a previous time interval as prediction. The first naïve approach uses the feature ‘previous week’, the second one the feature ‘previous day’, as prediction.

$$y(t) = y(t - t_{\Delta}) \quad (3)$$



For the third naïve model, we calculate the average for each interval per day of the training data. For classification, the target value is either 0 or 1. If the average value for the naïve model ‘average’ is below 0.5 for a specific time interval, the target value is set to 0. For regression, the naïve model ‘average’ gives the number of average occupied charging points for the specific time interval.

### 3.2.2. Linear Regression

The LinR model is a simple supervised learning algorithm that predicts continuous response variables. It captures the linear relationship between the independent variable  $x$  and the dependent variable  $y$ . The linear model  $f$  is a linear function of input  $x = (x_1, x_2, \dots, x_p) \in R^p$  [56]:

$$f(x) = w_0 + w_1x_1 + \dots + w_px_p \quad (4)$$

where  $w_0, w_1, \dots, w_p$  are weight parameters. The standard choice for the loss function is least squares, which is given by

$$L = \frac{1}{2} \sum_{i=1}^N (f(x_i) - y_i)^2 = \frac{1}{2} (XW - Y)^T (XW - Y) \quad (5)$$

where  $X \in R^{N \times (p+1)}$ ,  $W \in R^{p+1}$  and  $Y \in R^N$ . A regularization term is often added to the loss function to avoid overfitting. Common choices are  $L1$  or  $L2$  regularization. These regression models are called Lasso regression and Ridge regression, respectively. By taking the partial derivative of  $L$  with respect to  $W$ , we obtain the estimation  $W^*$  of the parameter  $W$  [56].

$$W^* = (X^T X)^{-1} X^T Y \quad (6)$$

LinR is a simple algorithm and is easy to interpret. However, it is sensitive to outliers and does not perform well if the relationship between independent and dependent variables is nonlinear. It also assumes that the input features are independent. Multicollinearity in input features should be carefully handled to obtain a reliable model [56].

### 3.2.3. Logistic Regression

LogR is a supervised learning algorithm which predicts a probability of given data being classified as a binary class 0 or 1 [57]. Despite its name, LogR is an algorithm for classification rather than regression. While a LinR model predicts a continuous outcome, a LogR model uses a function called sigmoid to scale the output of a linear model into a probability. The sigmoid function  $\sigma \rightarrow (0, 1)$ , which is also called logistic function, is defined as

$$\sigma(x) = \frac{1}{1 + \exp(-x)} \quad (7)$$

Given an output of a linear model  $z = w^T x + b$  where  $w$  is weight vector and  $b$  is bias, a sigmoid function  $\sigma$  is applied to  $z$ . Therefore, the probability of input  $x$  being classified as 1 is

$$P(y = 1|x) = \sigma(z) = \sigma(w^T x + b) = \frac{1}{1 + \exp(-(w^T x + b))} \quad (8)$$

### 3.2.4. K-Nearest Neighbors

K-nearest neighbors (KNN) is a nonparametric algorithm which does not require any assumption for the underlying data distribution [58]. KNN can be used for both classification and regression tasks. By using a local average, KNN regression predicts the value of the output variable. KNN classification aims to predict the class to which the output variable belongs by computing the local probability. Given an input data and a hyperparameter  $k$ , the model computes its class probability using class distribution of  $k$  nearest neighbors. Distance measures such as Euclidean distance or Mahalanobis distance

can be used to find nearest neighbors. The probability of an input data  $x$  belonging to class  $c$  is given by

$$p(y = c|x, k) = \frac{1}{k} \sum_{i \in N_k(x)} I(y_i = c) \quad (9)$$

where  $I$  is an indicator function which outputs 1 if the given condition is satisfied and 0 if otherwise. To achieve a good performance in KNN, it is important to find a good  $k$ . A small value of  $k$  means that noise will have a higher influence on the results of the classification, and respectively, regression. With larger value of  $k$ , noise can be reduced; however, the class prediction tends to be the majority class when the class distribution of the dataset is skewed. An optimal  $k$  can be chosen by hyperparameter-tuning methods such as cross-validation. Although KNN is a simple and intuitive algorithm, it has some limitations. KNN starts computation only with new input data, hence it is called a lazy learner. Additionally, the algorithm is not effective for high-dimensional data since the computational complexity grows exponentially as the number of dimensions increases [58].

### 3.2.5. Random Forest

Random forest (RF) is a supervised learning algorithm which uses an ensemble learning method for classification and regression [59]. Bagging, also known as bootstrap aggregating, is a method to train multiple classifiers on different data samples and decision-tree algorithms. The reason for this is to separate data in a way that achieves the largest information gain or reduces impurity in data. In RFs, the bagging method is used to train multiple decision trees. Data are randomly sampled with replacement from the entire dataset. Therefore, each decision tree has its own dataset for training. There is an additional bagging step called feature bagging. While features are selected by computing entropy or Gini index in decision trees, the features are randomly selected with feature bagging in RFs. Given  $n$  features, usually  $\sqrt{n}$  features are randomly chosen for classification tasks and  $n/3$  for regression tasks [59]. For one decision tree, multiple feature sets can be created and the best feature set could be chosen by cross validation using the dataset for the specific decision tree. The final prediction  $f$  using RF is the aggregation of outputs from all decision trees, given by

$$f = \frac{1}{B} \sum_{n=1}^B f_n(x) \quad (10)$$

where  $B$  is the number of decision trees,  $f_n$  is the output of the  $n$ th decision tree, and  $x$  is the unseen data.

RF shows good performance for unseen data, since it performs well at generalization and does not overfit. However, disadvantages of RFs are difficulties in interpretation and incremental learning [59].

### 3.2.6. XGBoost

XGBoost (eXtreme Gradient Boosting) is an ensemble model with boosting that can be used for regression and classification. Boosting is a method to train several classifiers in sequence by giving higher weight to hard samples which are misclassified by the previous classifier. The final prediction is given by a weighted sum of outputs of all the classifiers. There are different boosting algorithms such as AdaBoost and Gradient Boosting. XGBoost is an extension of Gradient Boosting which makes use of a gradient descent algorithm to minimize errors in the previous tree model [60]. XGBoost made improvements in speed and performance compared to Gradient Boosting. Parallelization of tree construction is a key feature of XGBoost which enables faster computation. XGBoost also prevents overfitting by introducing a regularization term in its objective function, and thus penalizes complex models. Other features such as tree pruning, cache, and sparse awareness facilitate efficient learning by XGBoost. XGBoost is one of the best tree-based models at present, thus it can be a good choice for many real-world tasks [60].

### 3.3. Evaluation Metrics

The following section describes the metrics used to evaluate the classification and regression models.

#### 3.3.1. Classification Metrics

One of the most common methods to evaluate classification models' performance is to create a confusion matrix. A confusion matrix is a table that shows the number of samples which are correctly classified or misclassified, as shown in Table 3 [61,62].

**Table 3.** Confusion matrix.

		Predicted	
		1	0
Actual	1	True Positive ( <i>TP</i> )	False Negative ( <i>FN</i> )
	0	False Positive ( <i>FP</i> )	True Negative ( <i>TN</i> )

Several evaluation metrics can be derived from the confusion matrix. The most widely used metric is accuracy, which is defined by

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN} \quad (11)$$

This is the proportion of correctly classified cases among all cases. However, accuracy is not appropriate if the data class distribution is skewed. Other metrics, such as recall and precision, can be used to evaluate performance of a model on imbalanced datasets. Recall is the proportion of actual positive cases that are correctly classified. It is also called sensitivity or true positive rate (TPR).

$$\text{Recall} = \frac{TP}{TP + FN} \quad (12)$$

This is useful for tasks such as cancer detection, i.e., when the cost of false-negative samples is high. Precision, which is also called positive predictive value (PPV), is the proportion of predicted positive cases that are actually positive.

$$\text{Precision} = \frac{TP}{TP + FP} \quad (13)$$

Precision is useful when the cost of false positive samples is high, for example, email spam detection. F1-score takes into account both precision and recall. It is defined by the harmonic mean of recall and precision.

$$F1 = 2 \times \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} \quad (14)$$

A receiver operating characteristic (ROC) curve shows the performance of a classification model at different thresholds. It is a plot of true positive rate (TPR), which is recall, and false-positive rate (FPR), which is 1-specificity. Specificity is the proportion of actual negative cases that are correctly classified to be negative. Thus, FPR is the proportion of predicted negative cases that are misclassified as negative cases. If a classifier has a threshold  $\tau = 1$ , it predicts all the samples as negative and has specificity = 1 and recall = 0. If a classifier has  $\tau = 0$ , it predicts all the samples as positive and has specificity = 0 and recall = 1. If the curve reaches the point (0, 1), which means specificity = 1 and recall = 1, it is a perfect classifier. The ROC curve of a random classifier is a diagonal line where TPR is equal to FPR.

The area under the curve (AUC) is a measure of model performance described by an ROC curve. The higher the AUC is, the better the model performance is. Since the range of an ROC curve is a unit square, a perfect classifier has an AUC of 1.

### 3.3.2. The Regression Metrics

LinR model tries to make the smallest difference between the predicted value and the observed value. The difference between these two values is called the residual. Therefore, the metric to evaluate the performance of the regression model uses the residual. The most common metric is root-mean-squared error (RMSE) [55]. RMSE is the square root of mean squared error, which is the average of squared residuals. It is defined by

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (15)$$

where  $y_i$  is the observed value of  $i$ th data and  $\hat{y}_i$  is the predicted value for  $y_i$ . The model fits better if RMSE is smaller.

R-squared ( $R^2$ ), also called the coefficient of determination, measures how much variance of the dependent variable is explained by the independent variables [63].  $R^2$  is given by

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (16)$$

The maximum value of  $R^2$  is 1, which means the model prediction is perfect.

## 4. Results

The following section highlights the results of the hyperparameter tuning for the different classification and regression models, and compares the obtained evaluation metrics for the different feature sets, modes and sites. The results were obtained from a single run of each algorithm with the respective optimal hyperparameter set.

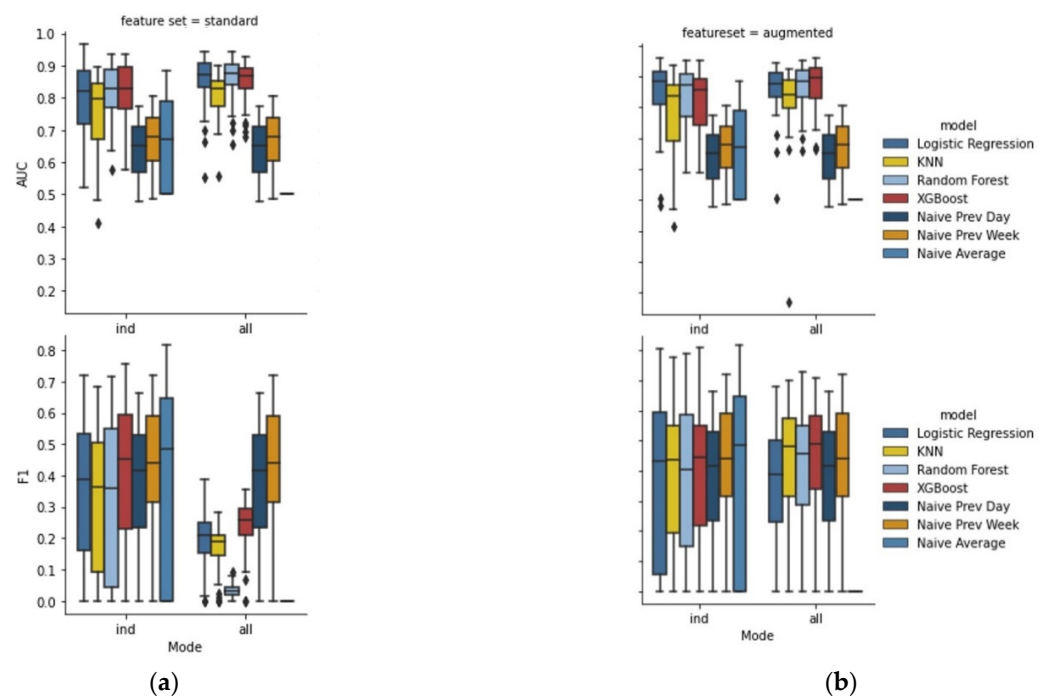
### 4.1. Results Classification

To obtain the optimal hyperparameters, we used a grid-search algorithm, choosing the hyperparameters where the model performs best on the validation set. The models were retrained before being validated against the validation set. Therefore, a different model was validated each time. After finding the optimal parameters, training was repeated on the entire training set with optimal parameters. In Table 4, an overview of the hyperparameter search for the classification task is given. For each parameter with a continuous value range, a linearly spaced vector of five elements was created as parameter grid. The four rightmost columns represent the corresponding values that performed best on the validation set when trained on all EVSE distinguished by standard and augmented feature sets as well as at site JPL and Caltech.

**Table 4.** Overview of classification hyperparameters used for grid search.

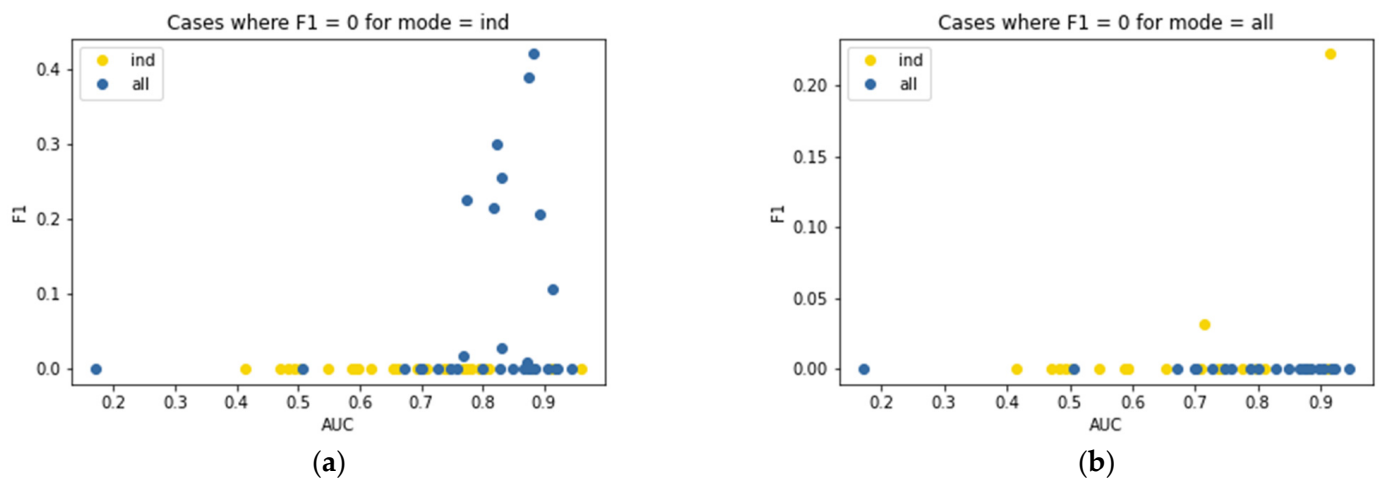
Model	Parameter	Values Range	Value Caltech (All, Stand.)	Value Caltech (All, Augm.)	Value JPL (All, Stand.)	Value JPL (All, Augm.)
KNN	n neighbors	[50, 300]	300	237	300	237
LogR	C	[0.1, 20]	0.1	0.1	10.005	5.075
RF	criterion	{gini, entropy}	entropy	gini	entropy	gini
	n estimators	[50, 300]	300	50	300	175
	max_depth	[2, 10]	10	10	10	8
XGBoost	n estimators	[50, 300]	300	50	50	50
	max_depth	[2, 10]	10	4	4	4

Figure 3 shows the distribution for the AUC (top) and F1 (bottom) scores of each charging point for the Caltech site. On the left-hand side, the standard feature set is used, and on the right-hand side, the augmented one. Each of the four models performed better on the AUC score than the naïve approaches. Using the augmented feature set, AUC score for both modes was increased. LogR in mode ‘individual’ benefits from the additional features. Overall XGBoost performed best on AUC score for mode ‘all’ on the augmented set with a median of 0.89, but LogR is only slightly worse for mode ‘individual’ with a median of 0.88. The application of the naïve approach ‘average’ for mode ‘all’ resulted in an AUC score of 0.5 and a F1 score of 0. This indicates that the model did not distinguish between class points. Rather, the naïve approach ‘average’ for mode ‘all’ predicts 0 for all data points. Considering median F1 score, XGBoost performs slightly better than the naïve approaches ‘previous week’ and ‘previous day’ on mode ‘individual’, whereas the naïve approach ‘average’ reaches the highest F1 score. By using the augmented feature set, an improvement of the F1 score is noticeable, especially for mode ‘all’ where the median for XGBoost increases by 0.23 and for RF by 0.42. While XGBoost has the highest median F1 score with 0.49, it is hardly higher than the one from the naïve approach ‘average’. It is noticeable that the mode ‘individual’ has a greater dispersion of  $F_1$  scores, but also higher values are achieved for individual charging points compared to mode ‘all’. Except for LogR, mode ‘all’ achieves generally slightly better median results for the augmented feature set at Caltech.



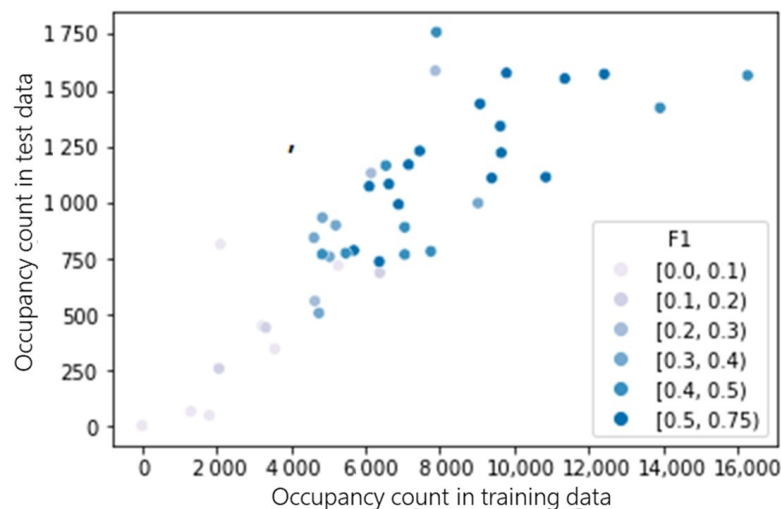
**Figure 3.** (a) AUC and F1 score for the different models and modes at Caltech for standard feature set; (b) AUC and F1 score for the different models and modes at Caltech for augmented feature set.

To further investigate which mode is favorable, we examined all predictions of the four models: LogR, KNN, RF and XGBoost with a F1 score equal to 0. On the left, Figure 4 shows the F1 and AUC score for all charging points with a F1 score of 0 for the mode ‘individual’ and the corresponding pair for the same charging point for mode ‘all’. On the right, the mode ‘all’ and F1 scores equal to 0 are displayed with the corresponding score pair for mode ‘individual’.



**Figure 4.** (a) F1 score equal to 0 for mode ‘individual’ and corresponding scores of mode ‘all’; (b) F1 score equal to 0 for mode ‘all’ and corresponding scores of mode ‘individual’.

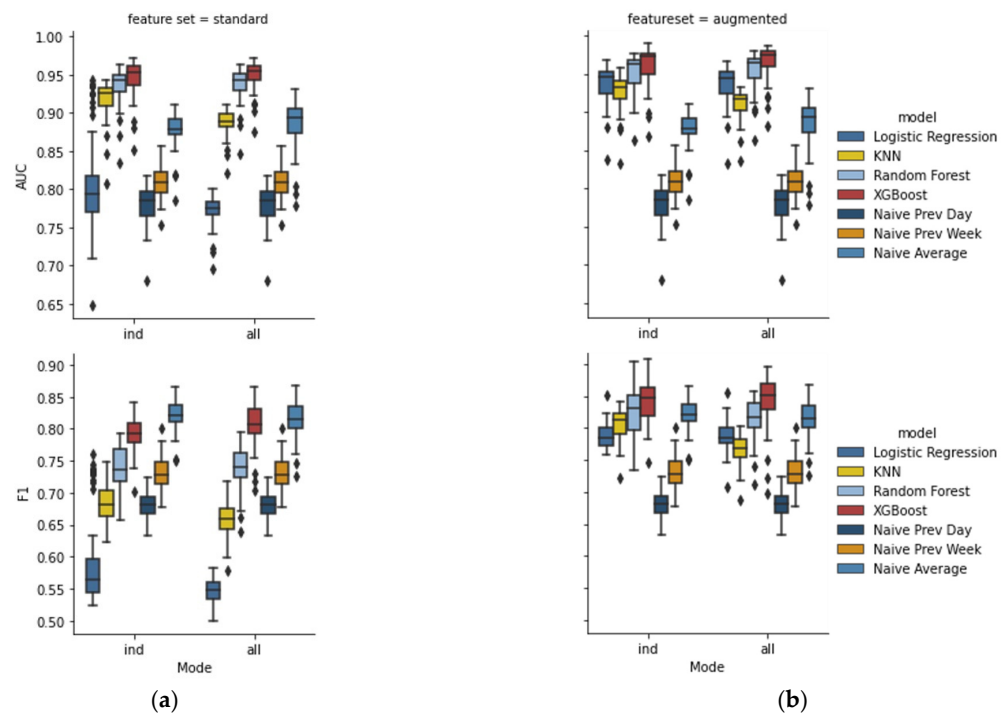
The left graph in Figure 4 displays few cases where using mode ‘all’ leads to F1 scores from around 0.1 up to 0.4 higher than the ones for mode ‘individual’. Comparing the two graphs shows that there are more model predictions where the mode ‘all’ achieves F1 scores higher than 0 than the other way round. To determine when F1 score is low, Figure 5 shows the datapoints of occupancy in the test data against the datapoints of occupancy in the training data, colored by their corresponding F1 score for mode ‘individual’.



**Figure 5.** Occupancy count in training and test data and corresponding F1 score.

The figure illustrates that charging points have a higher F1 score if they have higher occupancy and a lower one if they are rarely occupied. For charging points with a low relative occupancy, a model that has been trained on all charging points can therefore achieve better results in some cases. This is reflected in the lower scatter of the F1 score for the augmented feature set and mode ‘all’ compared to mode ‘individual’ in Figure 3.

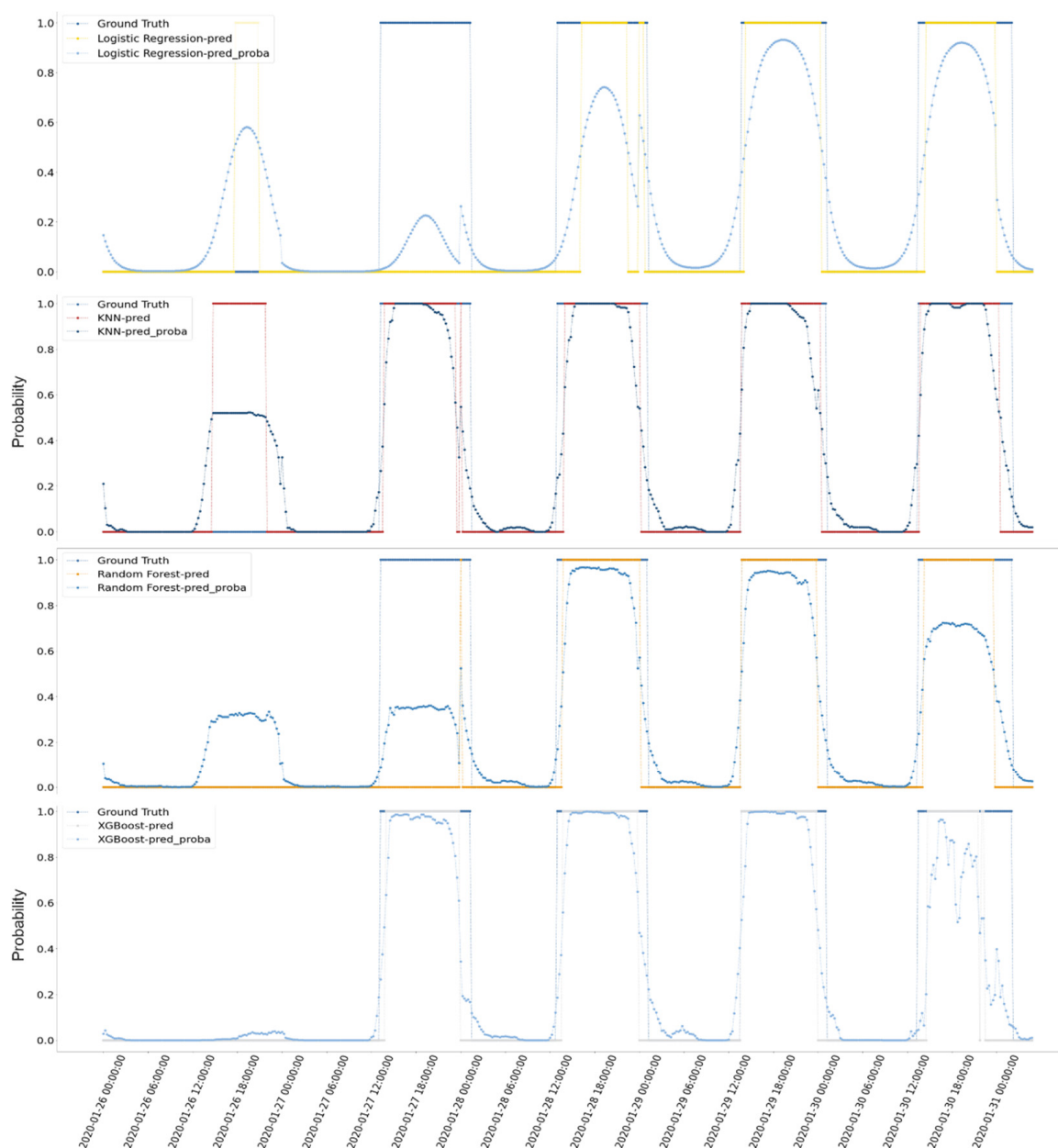
Figure 6 shows the distribution for the AUC (top) and F1 (bottom) scores for each feature set and mode for the JPL site.



**Figure 6.** (a) AUC and F1 score for the different models and modes at JPL for standard feature set; (b) AUC and F1 score for the different models and modes at JPL for augmented feature set.

Each of the four models performs better on the median AUC score than the naïve approaches, except for LogR and mode ‘individual’ as well as KNN and mode ‘all’. As with Caltech, the AUC score is noticeably higher when using the augmented feature set, and in total, XGBoost achieves the best median AUC score for mode ‘all’ on the augmented set. The advantage of the augmented feature set is particularly evident concerning the F1 score. For the standard feature set, none of the models perform better than the naïve approach ‘average’. Using the augmented feature set RF for mode ‘individual’ and XGBoost results in higher F1 scores than the naïve approach ‘average’, and XGBoost again obtains the highest score. Since JPL does not have any charging points with very low relative occupancies compared to Caltech, the comparison of the individual and all modes shows more similar performance metrics. Using the augmented feature set, the mode ‘individual’ achieves the best results for XGBoost in specific cases, evident from the longer boxplot whiskers, while the mode ‘all’ is slightly better when considering the median.

Figure 7 shows actual occupancy (ground truth), predicted probability, and resulting prediction for the four models for the last week of January 2020 and a sample charging point at the JPL site based on mode ‘all’ and standard feature set. If the ground truth is 1, the charging point is occupied, and available when 0. Considering the first day, 26 January 2020, LogR and KNN falsely predict the charging point will be occupied for a short amount of time in the afternoon, while RF and XGBoost forecasts are correct, with XGBoost’s probability being close to 0. This day was a Sunday, so the probability that a charging point at a workplace is occupied can be assumed to be near 0. On the second day, LogR and RF fail to predict the occupancy of the charging point, in contrast to KNN and XGBoost, which make a correct prediction for most of the day. All in all, for this sample week, XGBoost performs best, although the probability is worse on the last day compared to the other algorithms.



**Figure 7.** Forecasted occupancy probability for last week of January 2020 for one charging point at JPL based on mode ‘all’ and standard feature set.

The comparison of the two sites shows that the addition of features can improve performance, and that, in the case of Caltech for mode ‘individual’, even the comparatively simple linear model LogR performs better than the advanced models. In addition, for the augmented feature set, all four models outperform the naive approaches in terms of AUC score. For Caltech, the heterogeneous distribution of the average charging point occupancy is reflected in the scatter of the F1 score, whereby the mode ‘all’ leads to a reduction in this. With JPL, a narrower distribution can be seen and considerably better results are achieved overall. For the augmented feature set, the mode ‘individual’ achieves better results in the individual case both at Caltech and at JPL, while the mode ‘all’ performs better in terms of the median. For locations with a large discrepancy in the relative charging point occupancy, using a model trained on all data for charging points with a low relative occupancy can lead to better predictions. For certain charging points, an individually trained model achieves better performance. Overall, both approaches achieve very good results for the JPL site using the augmented feature set.



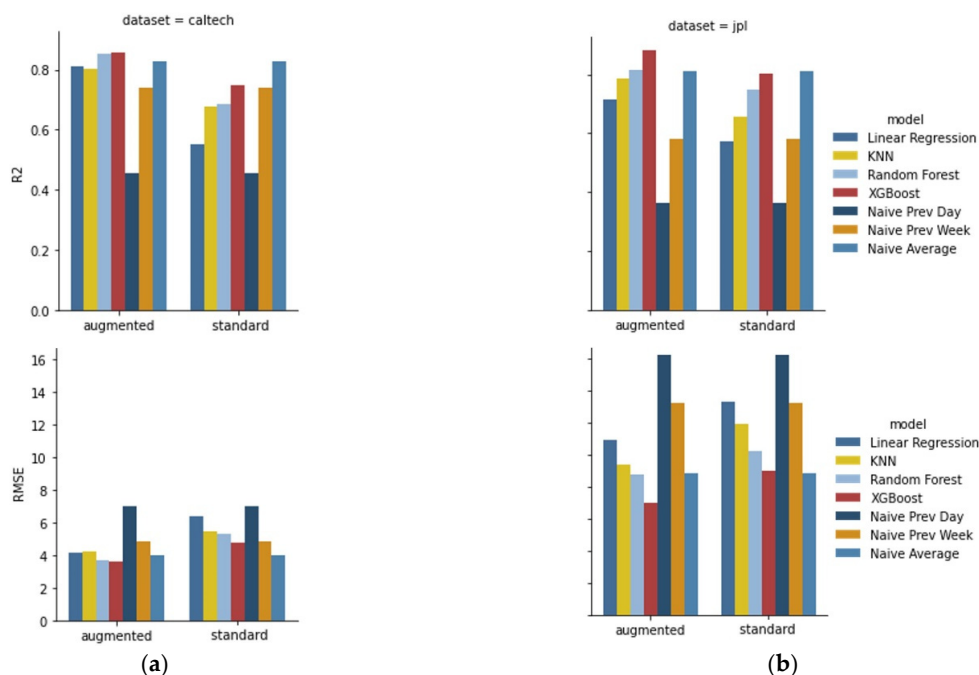
### 4.2. Results Regression

The following section shows the results of the hyperparameter optimization and discusses the obtained metrics of the different regression models. As in Section 4.1, for each parameter with a range of continuous values we used a linearly spaced vector of five elements for the parameter grid. Table 5 shows the corresponding values that performed best on the validation set distinguished by standard and augmented feature set and site JPL and Caltech.

**Table 5.** Overview of regression hyperparameters used for grid search.

Model	Parameter	Values Range	Value Caltech (All, Stand.)	Value Caltech (All, Augm.)	Value JPL (All, Stand.)	Value JPL (All, Augm.)
KNN	n_neighbors	[20, 200]	110	20	110	20
LinR	alpha	$[1 \times 10^{-5}, 1]$	1	$1 \times 10^{-5}$	$1 \times 10^{-5}$	$1 \times 10^{-5}$
RF	n_estimators	[50, 300]	112	175	50	175
	max_depth	[2, 10]	10	10	6	8
XGBoost	n_estimators	[50, 300]	50	300	112	300
	max_depth	[2, 10]	6	2	4	2

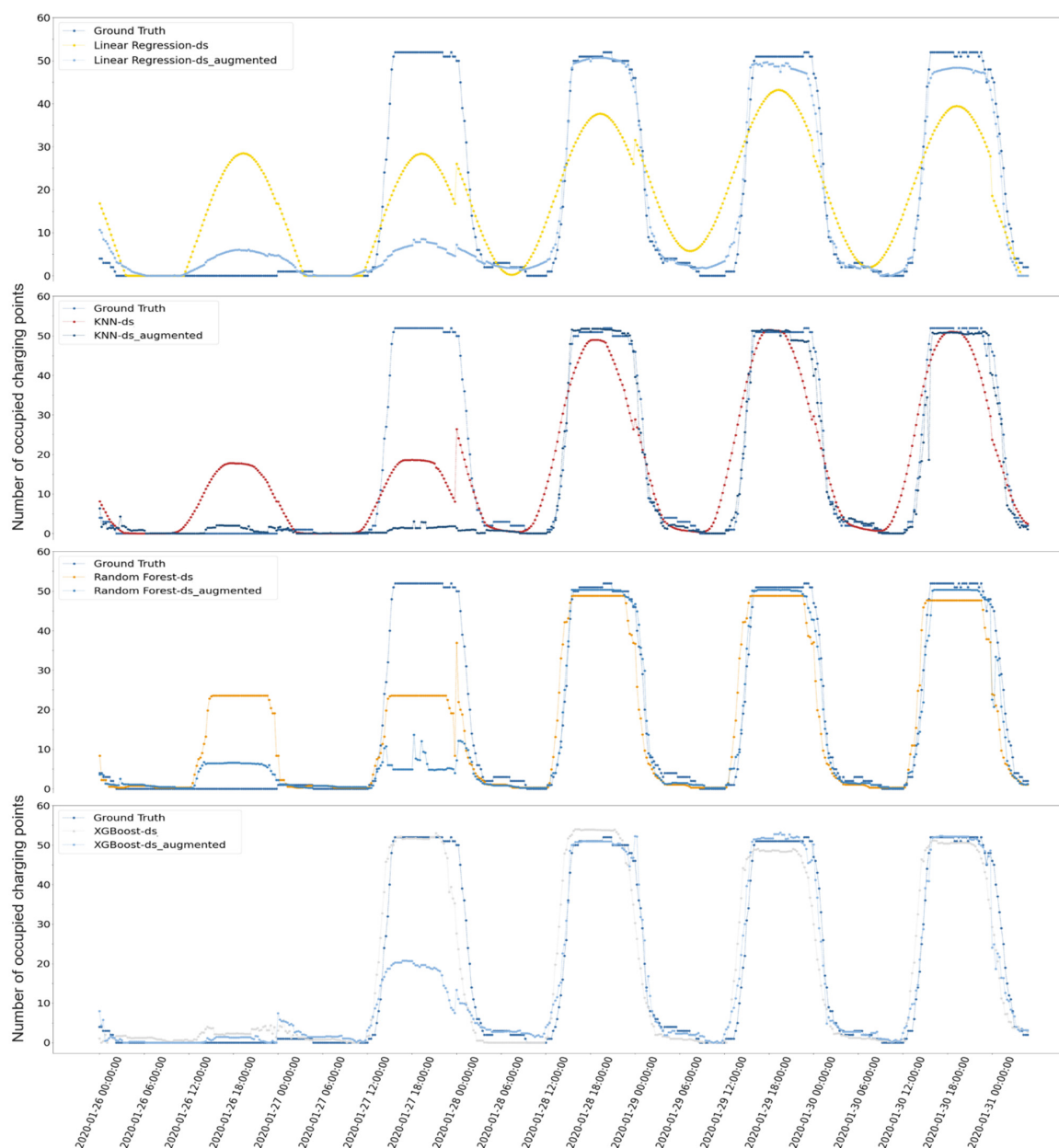
Figure 8 shows  $R^2$  (top) and RMSE (bottom) for all models differentiated by the standard and augmented feature set.



**Figure 8.** (a)  $R^2$  and RMSE for regression models at Caltech for standard and augmented feature set; (b)  $R^2$  and RMSE for regression models at JPL for standard and augmented feature set.

The left part represents Caltech, and on the right, JPL is displayed. The naïve approach ‘previous day’ performs worst in all cases. Considering Caltech and the standard feature set, only XGBoost has better scores than the naïve approach ‘previous week’, while the approach ‘average’ has the highest  $R^2$  score. The results of the naïve approach ‘average’ are similar to those achieved by RF for JPL and augmented feature set. Over both sites, XGBoost is better than the naïve approaches when using the augmented feature set, and using XGBoost gave the best scores both for  $R^2$  and RMSE metrics. Like the classification task, using additional features yields higher performance rates. In contrast to the classification task, the models

perform better for Caltech in regard to the RMSE. One of the reasons for this is the lower maximum occupancy number compared to JPL. Caltech has 54 charging points, but for the test set, a maximum occupancy number of 34 and average occupancy of 7, while JPL has a maximum of 52 out of 52 and an average of 14.4. Figure 9 displays the actual number of occupied charging points (ground truth) and the predicted number for both feature sets (standard = ds\_, augmented = ds\_augmented) for the four models for the last week of January 2020. KNN, for example, shows a noticeable deviation on Monday 27 January in relation to the augmented feature set forecast. The model treats the day similarly to the day before, which was not a working day. One possible reason why the model behaves this way is that Monday of the previous week (20 January) was a holiday in California, which is reflected in the augmented feature set. Looking at the results for KNN and the standard feature set, the prediction is not affected by this as much, but still shows a high deviation from the ground truth.



**Figure 9.** Forecasted occupied number of charging points for last week of January 2020 for JPL for both feature sets.

## 5. Discussion

Our results demonstrate that using supervised learning algorithms are an appropriate method to predict specific and overall charging station occupancy. The applied models were chosen in such a way that they represent a wide variety of supervised learning algorithms. LogR and LinR depict rather simple linear models, while KNN is considered a simple non-linear method. RF and XGBoost represent nonlinear tree-based ensemble learning methods. Additionally, XGboost is an advanced implementation of the gradient boosting algorithm and is widely used due to its tendency to yield highly accurate results. Regarding the limitations of using only these algorithms, it could be argued that using different models such as support vector machines or neural networks might achieve better results. We used three different naïve approaches, which provide a robust benchmark to compare the results of the applied supervised learning algorithms. Planned comparisons revealed that the naïve approach ‘average’ provides good results despite its simplicity. The analysis supports the findings of Motz et al., that adding features leads to substantially better results [49]. In line with the previously mentioned study, we find that higher relative charging occupancy of a charging point leads to better predictability of charging point occupancy [49]. Contrary to the findings of Motz et al., we did not find that using individual charging point data is generally superior to using all data, especially considering the F1 score [49]. By using the F1 score as optimization metric for the hyperparameter tuning, better results might be achieved. Another limitation of this work are ranges that were set for the hyperparameter tuning. Referring to the optimal hyperparameters presented in Tables 4 and 5 for the standard feature set, the parameters are sometimes at the limit of the grid search. Consequently, there may be room for improvement by extending these ranges. Another limitation is that we conducted a single run for each algorithm. Using different seeds and presenting the average of multiple runs might decrease performance. Since in some cases, e.g., for Caltech, the results are quite close, results from multiple runs might have led to the naïve approach ‘average’ exceeding the other algorithms. Considering the regression task, the naïve approach ‘average’ yields comparatively good results. One could argue that small improvements on the forecast do not match the additional overhead and complexity accompanied by the implementation of supervised learning algorithms such as XGBoost. For a real-world scenario, where an app or webpage would use the predictions to inform users on which charging point will be occupied, data are most likely received over time in streams of instances or batches. In this case, models use all the information to a specific time step  $t$  to predict the new target variable arriving at time step  $t + 1$ , which is known as incremental learning [64]. However, changes caused by dynamic environments or events such as COVID-19 bear the risk of concept drift. Concept drift means that the relationship between the input and target variable changes over time. Therefore, this evolution would most likely negatively impact the model’s prediction. To overcome concept drift, Hoens et al. discussed several options classified by three main categories: adaptive base learners, learners which modify the training set and ensemble techniques [64]. Zukov et al. further proposed a proximity-driven streaming random forest, which exploits combinations of the aforementioned [65]. Therefore, an apparent limitation of this study is that we used a static dataset for training and testing our models and excluded the COVID-19 period. It remains unclear how well the applied models could process incremental learning and the change in the underlying data.

## 6. Conclusions

This research applies four classification and four regression models to charging session data to predict charging point occupancy for a public and a workplace site using the ACN-data. As a benchmark, we compare our models to three naïve models with different metrics. Hyperparameter tuning was performed using a grid-search algorithm to find the optimal model configurations for both classification and regression using two different feature sets for both sites. Considering additional features improved the models’ predictions. Two training modes were compared for the classification task. Models were trained either on all

data, leading to a single model for each site, or on individual charging point data, resulting in one model per charging point. Overall, both approaches lead to high scores for the workplace site, while charging points with a higher relative occupancy rate had higher F1 scores. Which mode is more preferable depends on the relative charging point occupancy. All in all, using mode ‘all’ achieved slightly better results considering the median. For both classification and regression tasks, the model XGBoost showed the best results. Future work should consider implementing incremental learning and using all available data, including data from the COVID-19 period, to investigate the effects of concept drift. The possibility of comparing deep-learning methods with machine-learning algorithms to predict charging point occupancy warrants further investigation. Furthermore, future studies should aim to replicate results using different datasets. On top of that, integrating charging point occupancy prediction into an app or webpage is desirable in future work.

**Author Contributions:** Conceptualization, A.O. and Y.F.; methodology, A.O.; software, K.O. and Y.F.; validation, Y.F. and A.O.; formal analysis, A.O.; investigation, K.O. and Y.F.; data curation, K.O. and Y.F.; writing—original draft preparation, A.O., Y.F. and H.K.; writing—review and editing, A.O., Y.F. and H.K.; visualization, K.O., Y.F. and A.O.; project administration, A.O.; funding acquisition, A.O. All authors have read and agreed to the published version of the manuscript.

**Funding:** The data preparation and initial modelling approaches were conducted as part of the activities of the Forschungsstelle für Energiewirtschaft e.V. (FfE) in the project “Bidirektionales Lademanagement (BDL)”. The project is funded by the Federal Ministry for Economic Affairs and Climate Action (BMWK) (funding code: 01MV18004F). In-depth analyses of the paper were carried as part of the activities of FfE in the project “unIT-e<sup>2</sup>: Reallabor für verNETZtze E-Mobilität”. The project is funded by the Federal Ministry for Economic Affairs and Climate Action (BMWK) (funding code: 01MV21UN11).

**Acknowledgments:** Special thanks to Helena Dobes and Ryan Harper.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

ACN	Adaptive Charging
AUC	Area under the ROC Curve
EU	European Union
EV	Electric Vehicle
EVSE	Electric Vehicle Supply Equipment
FN	False Negative
FP	False Positive
FPR	False Positive Rate
GHG	Green House Gas
KNN	k-Nearest Neighbor
LinR	Linear Regression
LogR	Logistic Regression
RF	Random Forest
RMSE	Root Mean Square Error
ROC	Receiver Operating Characteristic
TN	True Negative
TP	True Positive
TPR	True Positive Rate
XGBoost	eXtreme Gradient Boosting

## References

1. Paris Agreement. *Ausgefertigt am 2015–12, Version vom 2015*; United Nations: Paris, France, 2015.
2. *EEA Greenhouse Gases—Data Viewer*; European Union: Brussels, Belgium, 2021; Available online: <https://www.eea.europa.eu/data-and-maps/data/data-viewers/greenhouse-gases-viewer> (accessed on 10 February 2022).

3. Mathiesen, B.V.; Lund, H.; Connolly, D.; Wenzel, H.; Østergaard, P.A.; Möller, B.; Nielsen, S.; Ridjan, I.; Karnøe, P.; Sperling, K. Smart Energy Systems for coherent 100% renewable energy and transport solutions. In *Applied Energy 145 (2015)*; Aalborg University: Copenhagen, Denmark, 2015. [CrossRef]
4. Mathiesen, B.V.; Connolly, D.; Lund, H.; Nielsen, M.P.; Schaltz, E.; Wenzel, H.; Bentsen, N.S.; Felby, C.; Kaspersen, P.; Ridjan, I. *CEESA 100% Renewable Energy Transport Scenarios towards 2050—Technical Background Report Part 2*; Aalborg University: Copenhagen, Denmark, 2015.
5. Duscha, V.; Wachsmuth, J.; Eckstein, J.; Pfluger, B. *GHG-Neutral EU2050—A Scenario of an EU with Net-Zero Greenhouse Gas Emissions and Its Implications*; Umweltbundesamt: Dessau-Roßlau, Germany, 2019.
6. *National Emissions Reported to the UNFCCC and to the EU Greenhouse Gas Monitoring Mechanism*; European Environment Agency (EEA): Copenhagen, Denmark, 2020; Available online: <https://www.eea.europa.eu/data-and-maps/data/national-emissions-reported-to-the-unfccc-and-to-the-eu-greenhouse-gas-monitoring-mechanism-16> (accessed on 27 February 2022).
7. Bieker, G. *A Global Comparison of the Life Cycle Greenhouse Gas Emissions of Combustion Engine and Electric Passenger Cars*; The International Council on Clean Transportation: Berlin, Germany, 2021.
8. *Infrastructure for Charging Electric Vehicles—More Charging Stations but Uneven Deployment Makes Travel across the EU Complicated*; European Court of Auditors: Luxembourg, 2021; Available online: [https://www.eea.europa.eu/Lists/ECADocuments/SR21\\_05/SR\\_Electrical\\_charging\\_infrastructure\\_EN.pdf](https://www.eea.europa.eu/Lists/ECADocuments/SR21_05/SR_Electrical_charging_infrastructure_EN.pdf) (accessed on 10 February 2022).
9. Dornoff, J.; Mock, P.; Baldino, C.; Bieker, G.; Diaz, S.; Miller, J.; Sen, A.; Tietge, U.; Wappelhorst, S. *Fit for 55: A Review and Evaluation of the European Commission Proposal for Amending the CO<sub>2</sub> Targets for New Cars and Vans*; The International Council on Clean Transportation: Berlin, Germany, 2021.
10. *Klimaschutzprogramm 2030 der Bundesregierung zur Umsetzung des Klimaschutzplans 2050*; Bundesregierung: Berlin, Germany, 2020.
11. *Directive 2014/94/EU of the European Parliament and of the Council on the Deployment of Alternative Fuels Infrastructure. Ausgefertigt am 2014-10-22, Version vom 2017-11-18*; European Parliament and the Council: Brussels, Belgium, 2017.
12. Coffman, M.; Bernstein, P.; Wee, S. Electric vehicles revisited: A review of factors that affect adoption. *Transp. Rev.* **2016**, *37*, 79–93. [CrossRef]
13. Liao, F.; Molin, E.; van Wee, B. Consumer preferences for electric vehicles: A literature review. *Transp. Rev.* **2017**, *37*, 252–275. [CrossRef]
14. Rezvani, Z.; Jansson, J.; Bodin, J. Advances in consumer electric vehicle adoption research: A review and research agenda. *Transp. Res. Part D Transp. Environ.* **2015**, *34*, 122–136. [CrossRef]
15. Sierzechula, W.; Bakker, S.; Maat, K.; van Wee, B. The influence of financial incentives and other socio-economic factors on electric vehicle adoption. *Energy Policy* **2014**, *68*, 183–194. [CrossRef]
16. Pevec, D.; Babic, J.; Carvalho, A.; Ghiassi-Farrokhfal, Y.; Ketter, W.; Podobnik, V. A survey-based assessment of how existing and potential electric vehicle owners perceive range anxiety. *J. Clean. Prod.* **2020**, *276*, 122779. [CrossRef]
17. Noel, L.; de Rubens, G.Z.; Sovacool, B.K.; Kester, J. Fear and loathing of electric vehicles: The reactionary rhetoric of range anxiety. *Energy Res. Soc. Sci.* **2018**, *48*, 96–107. [CrossRef]
18. Pevec, D.; Babic, J.; Carvalho, A.; Ghiassi-Farrokhfal, Y.; Ketter, W.; Podobnik, V. Electric Vehicle Range Anxiety: An Obstacle for the Personal Transportation (R)evolution? In Proceedings of the 2019 4th International Conference on Smart and Sustainable Technologies (SpliTech), Split, Croatia, 18–21 June 2019; University of Zagreb: Split, Croatia, 2019. [CrossRef]
19. *Chargehub.com*; Mogile Technologies: Pointe-Claire, QC, Canada, 2022; Available online: <https://chargehub.com/en/> (accessed on 10 February 2022).
20. Openchargemap. Open Source. 2022. Available online: <https://openchargemap.org/site> (accessed on 10 February 2022).
21. *Evnavigation*; GPS Tuner Systems KFT: Budapest, Hungary, 2022; Available online: <https://evnavigation.com/> (accessed on 10 February 2022).
22. *Go TO-U*; GO TO-U Inc.: Los Angeles, CA, USA, 2022; Available online: <https://go-tou.com/en/map> (accessed on 10 February 2022).
23. *E-Mobility: Only 1 in 9 Charging Points in EU Is Fast*; ACEA: Brussels, Belgium, 2021; Available online: <https://www.acea.auto/press-release/e-mobility-only-1-in-9-charging-points-in-eu-is-fast/> (accessed on 10 February 2022).
24. Tomasov, M.; Kajanova, M.; Bracinik, P.; Motyka, D. Overview of Battery Models for Sustainable Power and Transport Applications. *Transp. Res. Procedia* **2019**, *40*, 548–555. [CrossRef]
25. Tamilselvi, S.; Gunasundari, S.; Karuppiyah, N.; Razak RK, A.; Madhusudan, S.; Nagarajan, V.M.; Sathish, T.; Shamim, M.Z.M.; Saleel, C.A.; Afzal, A. Review on Battery Modelling Techniques. *Sustainability* **2021**, *13*, 10042. [CrossRef]
26. Wang, Y.; Tian, J.; Sun, Z.; Wang, L.; Xu, R.; Li, M.; Chen, Z. A comprehensive review of battery modeling and state estimation approaches for advanced battery management systems. *Renew. Sustain. Energy Rev.* **2020**, *131*, 110015. [CrossRef]
27. Zhang, C.; Li, K.; Mcloone, S.; Yang, Z. Battery modelling methods for electric vehicles—A review. In Proceedings of the 2014 European Control Conference (ECC), Strasbourg, France, 24–27 June 2014. [CrossRef]
28. Tremblay, O.; Dessaint, L.-A. Experimental Validation of a Battery Dynamic Model for EV Applications. *World Electr. Veh. J.* **2009**, *3*, 289–298. [CrossRef]
29. Zhang, Y.; Lyden, S.; de la Barra, B.L.; Haque, M.E. Optimization of Tremblay’s battery model parameters for plug-in hybrid electric vehicle applications. In Proceedings of the 2017 Australasian Universities Power Engineering Conference (AUPEC), Melbourne, Australia, 19–22 November 2017; pp. 1–6. [CrossRef]

30. Hu, Y.; de Callafon, R.A.; Tian, N.; Fang, H. Modeling of Lithium-ion Batteries via Tensor-Network-Based Volterra Model. *IFAC-PapersOnLine* **2021**, *54*, 509–515. [[CrossRef](#)]
31. Sidorov, D.; Muftahov, I.; Tomin, N.; Karamov, D.; Panasetsky, D.; Dreglea, A.; Liu, F.; Foley, A. A Dynamic Analysis of Energy Storage with Renewable and Diesel Generation using Volterra Equations. *IEEE Trans. Ind. Inform.* **2020**, *16*, 3451–3459. [[CrossRef](#)]
32. Sidorov, D.; Sidorov, D.; Zhukov, A.; Foley, A.; Tynda, A.; Muftahov, I.; Panasetsky, D.; Li, Y. Volterra Models in Load Leveling Problem. *E3S Web Conf.* **2018**, *69*, 01015. [[CrossRef](#)]
33. Pagany, R.; Camargo, L.R.; Dorner, W. A review of spatial localization methodologies for the electric vehicle charging infrastructure. *Int. J. Sustain. Transp.* **2018**, *13*, 433–449. [[CrossRef](#)]
34. Metais, M.O.; Jouini, O.; Perez, Y.; Berrada, J.; Suomalainen, E. Too much or not enough? Planning electric vehicle charging infrastructure: A review of modeling options. *Renew. Sustain. Energy Rev.* **2021**, *153*, 111719. [[CrossRef](#)]
35. Micari, S.; Polimeni, A.; Napoli, G.; Andaloro, L.; Antonucci, V. Electric vehicle charging infrastructure planning in a road network. *Renew. Sustain. Energy Rev.* **2017**, *80*, 98–108. [[CrossRef](#)]
36. Guo, S.; Zhao, H. Optimal site selection of electric vehicle charging station by using fuzzy TOPSIS based on sustainability perspective. *Appl. Energy* **2015**, *158*, 390–402. [[CrossRef](#)]
37. Viswanathan, S.; Appel, J.; Chang, L.; Man, I.V.; Saba, R.; Gamel, A. Development of an assessment model for predicting public electric vehicle charging stations. *Eur. Transp. Res. Rev.* **2018**, *10*, 54. [[CrossRef](#)]
38. Flammini, M.G.; Prettico, G.; Julea, A.; Fulli, G.; Mazza, A.; Chicco, G. Statistical characterisation of the real transaction data gathered from electric vehicle charging stations. *Electr. Power Syst. Res.* **2019**, *166*, 136–150. [[CrossRef](#)]
39. Hecht, C.; Das, S.; Bussar, C.; Sauer, D.U. Representative, empirical, real-world charging station usage characteristics and data in Germany. *eTransportation* **2020**, *6*, 10007. [[CrossRef](#)]
40. Almaghrebi, A.; Aljuheshi, F.; Rafea, M.; James, K.; Alahmad, M. Data-Driven Charging Demand Prediction at Public Charging Stations Using Supervised Machine Learning Regression Methods. *Energies* **2020**, *13*, 4231. [[CrossRef](#)]
41. Xiong, Y.; Wang, B.; Chu, C.C.; Gadh, R. Electric Vehicle Driver Clustering using Statistical Model and Machine Learning. In Proceedings of the 2018 IEEE Power & Energy Society General Meeting (PESGM), Portland, OR, USA, 5–10 August 2018; University of California: Los Angeles, CA, USA, 2018. [[CrossRef](#)]
42. Lucas, A.; Barranco, R.; Refa, N. EV Idle Time Estimation on Charging Infrastructure, Comparing Supervised Machine Learning Regressions. *Energies* **2019**, *12*, 269. [[CrossRef](#)]
43. Sørensen, Å.L.; Lindberg, K.B.; Sartori, I.; Andresen, I. Analysis of residential EV energy flexibility potential based on real-world charging reports and smart meter data. *Energy Build.* **2021**, *241*, 110923. [[CrossRef](#)]
44. Gerritsma, M.; AlSkaiif, T.A.; Fidler, H.A.; van Sark, W.G. Flexibility of Electric Vehicle Demand: Analysis of Measured Charging Data and Simulation for the Future. *World Electr. Veh. J.* **2019**, *10*, 14. [[CrossRef](#)]
45. Sadeghianpourhamami, N.; Refa, N.; Strobbe, M.; Develder, C. Quantitative analysis of electric vehicle flexibility: A data-driven approach. *Int. J. Electr. Power Energy Syst.* **2018**, *95*, 451–462. [[CrossRef](#)]
46. Bikcora, C.; Refa, N.; Verheijen, L.; Weiland, S. Prediction of availability and charging rate at charging stations for electric vehicles. In Proceedings of the 2016 International Conference on Probabilistic Methods Applied to Power Systems (PMAPS), Beijing, China, 16–20 October 2016; Eindhoven University of Technology: Eindhoven, The Netherlands, 2016. [[CrossRef](#)]
47. Tian, Z.; Jung, T.; Wang, Y.; Zhang, F.; Tu, L.; Xu, C.; Tian, C.; Li, X. Real-Time Charging Station Recommendation System for Electric-Vehicle Taxis. *IEEE Trans. Intell. Transp. Syst.* **2016**, *17*, 3098–3109. [[CrossRef](#)]
48. Majidpour, M. *Time Series Prediction for Electric Vehicle Charging Load and Solar Power Generation in the Context of Smart Grid*; University of California: Los Angeles, CA, USA, 2016.
49. Motz, M.; Huber, J.; Weinhardt, C. Forecasting BEV charging station occupancy at workplaces. In *INFORMATIK 2020*; Reussner, R.H., Koziulek, A., Heinrich, R., Eds.; Gesellschaft für Informatik: Bonn, Germany, 2021. [[CrossRef](#)]
50. Lee, Z.J.; Li, T.; Low, S.H. ACN-Data: Analysis and Applications of an Open EV Charging Dataset. In Proceedings of the E-Energy '19: Tenth ACM International Conference on Future Energy Systems, Phoenix, AZ, USA, 25–28 June 2019; California Institute of Technology: Pasadena, CA, USA, 2019. [[CrossRef](#)]
51. Lee, G.; Lee, T.; Low, Z.; Low, S.H.; Ortega, C. Adaptive charging network for electric vehicles. In Proceedings of the 2016 IEEE Global Conference on Signal and Information, Washington, DC, USA, 7–9 December 2016; California Institute of Technology: Pasadena, CA, USA, 2016. [[CrossRef](#)]
52. Lee, Z.J.; Chang, D.; Jin, C.; Lee, G.S.; Lee, R.; Lee, T.; Low, S.H. Large-Scale Adaptive Electric Vehicle Charging. In Proceedings of the 2018 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids, Aalborg, Denmark, 29–31 October 2018; California Institute of Technology: Pasadena, CA, USA, 2018. [[CrossRef](#)]
53. Mahajan, T.; Singh, G.; Bruns, G.; Bruns, G.; Mahajan, T.; Singh, G. An Experimental Assessment of Treatments for Cyclical Data. In Proceedings of the 2021 Computer Science Conference for CSU Undergraduates, Virtual, 6 March 2021; Available online: <https://cscsu-conference.github.io/index.html> (accessed on 26 February 2022).
54. London, I. *Encoding Cyclical Continuous Features—24-h Time*; Ian London: New York, NY, USA, 2016; Available online: <https://ianlondon.github.io/blog/encoding-cyclical-features-24hour-time/> (accessed on 11 February 2022).
55. Gareth, J.; Daniela, W.; Trevor, H.; Robert, T. *An Introduction to Statistical Learning—With Applications in R*; Springer: New York, NY, USA, 2017.
56. Rebalá, G.; Ravi, A.; Churiwala, S. Regressions. In *An Introduction to Machine Learning*; OpsMx Inc.: San Ramon, CA, USA, 2019.

57. Bishop, C.M. *Pattern Recognition and Machine Learning*; Springer: New York, NY, USA, 2006.
58. Rebala, G.; Ravi, A.; Churiwala, S. Clustering. In *An Introduction to Machine Learning*; OpsMx Inc.: San Ramon, CA, USA, 2019.
59. Rebala, G.; Ravi, A.; Churiwala, S. Random Forests. In *An Introduction to Machine Learning*; OpsMx Inc.: San Ramon, CA, USA, 2019.
60. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. In Proceedings of the KDD '16: 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; Association for Computing Machinery: New York, NY, USA, 2016.
61. Murphy, K.P. *Machine Learning: A Probabilistic Perspective (Adaptive Computation and Machine Learning)*; MIT Press: Cambridge, MA, USA, 2012.
62. Murphy, K.P. *Machine Learning—A Probabilistic Perspective*; Massachusetts Institute of Technology: Cambridge, MA, USA, 2012.
63. Chicco, D.; Warrens, M.J.; Jurman, G. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Comput. Sci.* **2021**, *7*, e623. [[CrossRef](#)]
64. Hoens, R.T.; Polikar, R.; Chawla, N.V. Learning from streaming data with concept drift and imbalance: An overview. *Prog. Artif. Intell.* **2012**, *1*, 89–101. [[CrossRef](#)]
65. Zikov, A.V.; Sidorov, D.N.; Foley, A.M. Random Forest Based Approach for Concept Drift Handling. In *Analysis of Images, Social Networks and Texts. AIST 2016*; Communications in Computer and Information Science; Springer: Cham, Switzerland, 2016; Volume 661. [[CrossRef](#)]