

## Secondary Publication



NachnOstadi, Ali; Das, Debasree; Elmamooz, Golnaz; Nicklas, Daniela

## Data Engineering for Horizon-Aware Smart Parking Systems

Date of secondary publication: 21.09.2026

Version of Record (Published Version), Article

Persistent identifier: urn:nbn:de:bvb:473-irb-117289x

### Primary publication

Ostadi, Ali; Das, Debasree; Elmamooz, Golnaz; Nicklas, Daniela (2026): Data Engineering for Horizon-Aware Smart Parking Systems, in: Datenbank-Spektrum : Zeitschrift für Datenbanktechnologie ; Organ der Fachgruppe Datenbanken der Gesellschaft für Informatik e.V., Berlin ; Heidelberg: Springer, Vol. 26, No. 2, pp. 155–164, doi: 10.1007/s13222-026-00544-w.

### Legal Notice

This work is protected by copyright and/or the indication of a licence. You are free to use this work in any way permitted by the copyright and/or the licence that applies to your usage. For other uses, you must obtain permission from the rights-holders.

This document is made available under a Creative Commons license.



The license information is available online:

<https://creativecommons.org/licenses/by/4.0/legalcode>



# Data Engineering for Horizon-Aware Smart Parking Systems

Ali Ostadi<sup>1</sup> · Debasree Das<sup>1</sup> · Elmamooz Golnaz<sup>1</sup> · Daniela Nicklas<sup>1</sup> 

Received: 15 February 2026 / Accepted: 15 June 2026 / Published online: 16 July 2026  
© The Author(s) 2026

## Abstract

Accurate short-term parking availability forecasting is an important component of intelligent transportation systems. While numerous machine learning and deep learning models have been proposed, most approaches rely on a single predictor and implicitly assume uniform performance across forecast horizons, despite evidence that accuracy varies with lead time. This paper presents a systematic evaluation of multi-step (1–24 h) parking availability forecasting using sensor-based occupancy data from two urban parking facilities. We show that different model classes exhibit complementary strengths at different horizons, with no single model consistently achieving the lowest error across the full forecast range. To address this limitation, we propose a horizon-aware hybrid forecasting approach that combines two complementary base models using validation-driven weighted blending. Two variants are considered: a Global Weighted Blend with a shared weight and a Local Weighted Blend with horizon-specific weights.

Experimental results show that both hybrid approaches outperform individual baseline models, with the Local Weighted Blend achieving the best overall performance across all evaluation metrics. These findings demonstrate that horizon-aware blending provides a simple and effective strategy for improving multi-step parking availability forecasts. As an outlook, we propose a context-aware, adaptive model-management architecture for sensor-based smart city systems.

**Keywords** Parking availability forecasting · Multi-step time series forecasting · Horizon-aware forecasting · Machine learning · Deep learning · Decision support systems

## 1 Introduction

Since the 1990, the term smart city [1] is used to describe technology-induced urban transformation projects in various application domains, such as mobility, environmental sustainability, or local culture. Many rely on sensor data to obtain up-to-date real-world information, which are processed to derive actions and monitor their results, and thus can be seen as instances of a smart system cycle [2].

Nowadays, this processing often relies on machine learning models that are trained on historical sensor data and available context information [3]. However, when sensor installations are new, there are only few data available for training, which changes over time [4]; further more, the downstream tasks might be manifold, depending on the context, so that one trained and deployed model is not enough for operation. Instead, we propose architectures that use a multitude of models, adaptive model management, and context-aware selection and re-training when needed. In this paper, we illustrate this need by studying a common use case: smart parking systems, in which sensors are used to observe the current occupation of parking lots. By training machine learning models, this data can be used to predict future availability over different prediction horizons, where a horizon refers to the length of the future time window for which a prediction is made. For example, this may correspond to short-term horizons, such as one hour when driving towards the city center, or longer-term horizons, such as 24 h when planning a next-day trip. More generally, the approach can be applied to other sensor-based predic-

---

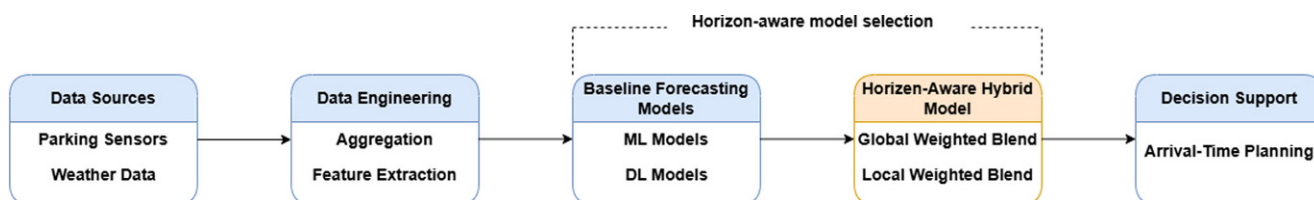
✉ Daniela Nicklas  
daniela.nicklas@uni-bamberg.de

Ali Ostadi  
ali.ostadiy@gmail.com

Debasree Das  
debasree.das@uni-bamberg.de

Elmamooz Golnaz  
golnaz.elmamooz@uni-bamberg.de

<sup>1</sup> Chair of Mobile Systems, University of Bamberg, Bamberg, Germany



**Fig. 1** Workflow of the proposed horizon-aware parking availability forecasting approach. Solid arrows represent the processing pipeline from data sources to decision support. The dotted line indicates the horizon-aware model selection process used to identify and combine complementary models for the hybrid forecasts

tion tasks with variable prediction horizons, such as taxi-demand prediction [5] or crowd management [2].

Finding available parking remains a persistent source of urban traffic congestion, resulting in increased travel time, fuel consumption, and emissions [6, 7]. In addition to spatial uncertainty, drivers face temporal uncertainty, as information on how parking availability will evolve in the coming hours is typically unavailable [6, 8]. While prior studies show that parking information can reduce cruising time, most existing systems focus on current occupancy rather than future availability [9, 10]. Consequently, enabling short-term parking availability prediction is a key challenge for improving arrival-time decisions and reducing inefficient parking search [6, 11]. Advances in sensing technologies enable continuous collection of parking occupancy data, forming the basis for data-driven prediction. Parking sensors record time-stamped occupied and free states, capturing daily, weekly, and event-driven demand patterns [6, 12]. Prior work shows that such occupancy time series are well suited for short-term forecasting due to strong temporal regularities [10, 13]. For drivers, forecasts over the next hours—particularly within a 24-hour horizon—are more informative than instantaneous availability, as they support planning based on anticipated demand rather than current conditions alone [6, 8]. As a result, sensor-based parking availability prediction has become a plausible method for improving parking information systems from a driver-oriented perspective [12, 14].

Many modeling approaches have been explored for predicting parking occupancy from sensor-based time series, ranging from statistical and regression-based methods using temporal covariates [6, 14] over classic machine learning models that capture nonlinear relationships in parking demand [11, 12] to deep learning methods that learn temporal dependencies directly from sequential inputs [13, 15]. However, comparative studies consistently show that higher model complexity does not guarantee superior accuracy across all horizons, and simpler models may outperform deep architectures [10, 14]. This motivates moving beyond one-model-fits-all forecasting toward horizon-aware modeling strategies.

Despite progress in parking occupancy prediction, most existing work treats forecasting as a single-model task, de-

ploying one algorithm across the full multi-step horizon [6, 11]. This assumption persists even though empirical evidence shows that model performance varies with forecast horizon [11, 14]. Moreover, while parking prediction is often motivated by mobility and driver-assistance benefits, fewer contributions demonstrate how multi-horizon forecasts are designed and evaluated in a way that supports arrival-time decision making [6, 8, 12]. To address these gaps, this work proposes a horizon-aware hybrid forecasting approach that combines complementary models through global and per-horizon weighted blending, as illustrated in Fig. 1.

The workflow begins with sensor-based parking and weather data, which are processed through aggregation and feature extraction to construct the modeling dataset. Based on this data, multiple baseline models from classic machine learning and deep learning are trained to generate multi-step (1–24 h) forecasts. Building on this perspective, we treat multi-horizon forecasting as a setting where model accuracy varies across horizons. Accordingly, the proposed approach incorporates a horizon-aware model selection process, where models are evaluated across forecast horizons to identify complementary strengths. Using sensor-based occupancy time series, we evaluate classic machine learning and deep learning models under a unified experimental protocol to ensure fair comparison [6, 14].

Beyond single-model selection, we introduce two horizon-aware blending strategies: a Global Weighted Blend with a single optimized weight across horizons, and a Local Weighted Blend with horizon-specific weights to exploit complementary strengths across the 24-step prediction window. The resulting hybrid forecasts are used to support arrival-time planning.

The contributions of this work are threefold: First, we provide a systematic comparison of statistical, machine learning, and deep learning models for short-term parking availability prediction using sensor-based data, with selection grounded in empirical performance rather than assumed complexity. Second, we propose a horizon-aware hybrid forecasting strategy via global and local weighted blending to address horizon-dependent performance variation. Third, we analyze short-term forecasts over a 24-hour horizon with a focus on supporting arrival-time planning,

thereby linking multi-horizon prediction to practical driver-oriented decision needs. In our outlook, we propose to use an adaptive architecture to disburden the data engineer from manual selection and deployment of ML models in a changing environment.

## 2 Related Work

Prior work on parking availability forecasting spans a wide range of data sources, prediction targets, and modeling techniques. This section briefly focuses on three aspects most relevant to this study: sensor-based forecasting, applied machine learning models, and horizon-dependent performance behavior.

For *sensor-based parking forecasting*, many studies rely on sensor-derived parking data, either in the form of event-based arrival and departure logs [11] or aggregated occupancy time series collected from parking facilities [14]. Historical availability is consistently identified as the dominant predictive signal, while contextual variables such as traffic conditions and weather are often included to enhance performance, with traffic-related features typically found to be more influential than weather [16]. High-resolution sensor data allow models to learn detailed short-term parking dynamics from recent history, while public datasets often report availability as percentages and rely on daily and weekly time patterns such as hour of day or day of week [17, 18].

For aspects of *applied machine learning and deep learning models*, a broad spectrum of predictive models has been evaluated for parking availability forecasting. Classic machine learning approaches, including decision trees, random forests, gradient boosting, and multilayer perceptrons, are widely reported as strong and robust baselines under real-world conditions [11, 19, 20]. Statistical and kernel-based methods such as ARIMA and support vector regression are also commonly studied, with SVR often outperforming purely linear models [14, 20]. More recently, deep learning architectures, predominantly recurrent and convolutional-recurrent models capture temporal dependencies directly from sequential inputs [13, 18]. Hybrid approaches combining neural networks with uncertainty-aware components, such as neuro-fuzzy systems, have also been proposed to improve robustness [15]. Overall, empirical results suggest that model effectiveness is highly context-dependent and not strictly correlated with architectural complexity [11, 17, 21].

*Horizon-dependent performance behavior:* An important challenge in multi-step parking availability forecasting is that prediction accuracy varies with forecast horizon. Several studies report increasing error for longer lead times and show that different model families dominate at different horizons [17, 22]. Comparative benchmarks furthermore

indicate that the best-performing model can change across datasets and horizons [18, 20]. Despite this, most existing approaches deploy a single model or a fixed hybrid across the full forecast range, without explicitly adapting model contribution to horizon-specific behavior [15, 16]. This gap motivates the horizon-aware hybrid forecasting strategy proposed in this work.

## 3 Data Preprocessing

This section describes the data sources and pre-processing steps used to construct the supervised learning dataset for multi-horizon parking availability forecasting.

### 3.1 Data Sources

Parking occupancy data were obtained from the project partner Smart City System Parking Solutions GmbH via the ParkingPilot API, which provides historical parking events in CSV format. The study considers two off-street parking facilities in Bamberg, Germany: *P+R Brennerstraße* (ID 38) and *P+R Heinrichsdamm* (ID 634).

*P+R Brennerstraße* is located near the central train station and has a total capacity of 118 spaces, including 5 reserved spaces. *P+R Heinrichsdamm* has a total capacity of 800 spaces; however, only 72 sensor-equipped outdoor spaces (including 5 reserved spaces) are considered in this study. For both facilities, the API provides space-level arrival and departure timestamps.

The dataset covers the period from 30 November 2024 to 28 November 2025, capturing daily, weekly, and seasonal demand patterns. All timestamps are provided in a consistent time reference. To account for external influences, hourly weather data (temperature, relative humidity, precipitation) were retrieved from the Open-Meteo historical weather service and temporally aligned with the parking data. All data are ingested automatically and stored in a centralized relational database, while raw observations are preserved to ensure reproducibility.

### 3.2 Preprocessing and Feature Construction

Raw parking event data are transformed into structured time series suitable for multi-step forecasting. For each parking lot  $l$ , minute-level occupancy is derived from arrival and departure intervals:

$$\text{occupied}_l(t) = \sum_{i=1}^{N_l} \mathbb{I}(a_i \leq t < d_i), \quad (1)$$

$$\text{free}_l(t) = C_l - \text{occupied}_l(t), \quad (2)$$

where  $a_i$  and  $d_i$  denote arrival and departure times,  $C_i$  is the sensor-covered capacity, and  $\mathbb{I}(\cdot)$  is the indicator function.

To reduce noise and align with operational forecasting needs, availability is aggregated to hourly resolution. For each hour, summary statistics of free spaces (mean, minimum, maximum, and range) are computed. Temporal features capturing periodic demand patterns (hour of day, day of week, calendar week, and month) are extracted.

Hourly parking availability is combined with meteorological variables. Short-term dynamics are represented using rolling statistics and first-order differences, including 24-hour rolling mean, standard deviation, and hour-to-hour changes. Categorical context variables (e.g., week-day/weekend and time-of-day regimes) are numerically encoded. The final feature set consists of 20 variables, with the target defined as the hourly average number of free parking spaces.

To prevent temporal leakage, data are split chronologically into 70% training, 15% validation, and 15% testing sets. A sliding-window approach is applied, where each input consists of the previous 48 h and the prediction target corresponds to the subsequent 24-hour trajectory:

$$\mathbf{X}^{(k)} \in \mathbb{R}^{48 \times p}, \quad \mathbf{y}^{(k)} \in \mathbb{R}^{24}, \quad p = 20. \quad (3)$$

Classic models use flattened inputs, while deep learning models preserve temporal structure. All numerical features are normalized using min-max scaling.

## 4 Method

This section presents the predictive models and training setup, followed by the horizon-aware weighted blending approach used for multi-step parking availability forecasting.

### 4.1 Predictive Models and Training Setup

To evaluate the performance of different models, we trained a diverse set of classic machine learning and deep learning models to forecast next-day (24-hour) parking availability. All models were trained on the same supervised dataset and evaluated under a unified experimental protocol to ensure comparability across parking facilities.

We selected hyperparameters based on prior literature focused on model architectures [23–25] and validation-based tuning, with the goal of balancing predictive accuracy and robustness across forecast horizons. The complete configuration of all models is reported in Table 1.

This table provides an overview of all evaluated models and their key hyperparameters. The table includes classic machine learning models, deep learning architectures, and

**Table 1** Predictive models and key hyperparameters.

| Model                     | Key hyperparameters                                       |
|---------------------------|-----------------------------------------------------------|
| Linear/Ridge Regression   | Multi-output regression; Ridge: $\alpha = 1.0$            |
| Decision Tree             | Max depth=20; min samples split=10; min leaf=5            |
| Random Forest             | 100 trees; max depth=20; min samples split=10; min leaf=5 |
| Gradient Boosting         | 50 estimators; depth=5; learning rate=0.1                 |
| XGBoost                   | 100 estimators; depth=5; learning rate=0.1                |
| SVR (RBF)                 | $C=1.0$ ; $\epsilon=0.1$ ; max 5000 samples               |
| RNN-based models          | LSTM/GRU/BiLSTM: two layers (150, 75 units)               |
| Seq2Seq (Encoder-Decoder) | LSTM encoder-decoder; 24-step output horizon              |
| CNN-RNN hybrids           | 1D CNN (64 filters, $k=3$ ) + RNN layers (150, 75)        |

hybrid configurations considered in this study. For each model, the most relevant hyperparameters and architectural settings are reported to ensure transparency and reproducibility of the experimental setup. The selected configurations reflect commonly used values in prior literature as well as validation-based tuning, aiming to balance model complexity, predictive accuracy, and robustness across different forecast horizons.

The evaluated classic machine learning baselines include Linear Regression, Ridge Regression [26], Decision Trees [27], Random Forests [28], Gradient Boosting [29], XGBoost [24], and Support Vector Regression with an RBF kernel [30]. Due to its computational cost, SVR was trained on a randomly selected subset of 5000 samples. Classic models are formulated as multi-output regressors, mapping flattened input features  $X \in \mathbb{R}^{N \times d}$  to a 24-step output vector  $Y \in \mathbb{R}^{N \times 24}$ .

Deep learning models operate directly on sequential inputs of the previous 48 h and generate a 24-hour forecast. Evaluated architectures include recurrent neural networks such as LSTM [31], GRU [32], and bidirectional LSTM [33], as well as an encoder-decoder (sequence-to-sequence) LSTM for direct multi-step prediction [34]. In addition, convolutional-recurrent hybrids (CNN-LSTM, CNN-GRU, CNN-Bidirectional LSTM) were considered, combining convolutional layers for short-term pattern extraction with recurrent layers for modeling longer-term temporal dependencies [35].

### 4.2 Horizon-Aware Weighted Blending

Our model comparison revealed that no single model consistently minimizes error across the full 24-hour forecast horizon. Some models perform best at short horizons, while others dominate at longer horizons, indicating that single-

model selection is suboptimal. To address this, we adopt a forecast combination strategy that blends predictions from two complementary base models. Let  $\hat{y}_{t,h}^{(A)}$  and  $\hat{y}_{t,h}^{(B)}$  denote predictions from models  $A$  and  $B$  at horizon  $h$ . The hybrid forecast is defined as:

$$\hat{y}_{t,h}^{(H)} = \alpha_h \hat{y}_{t,h}^{(A)} + (1 - \alpha_h) \hat{y}_{t,h}^{(B)}, \quad \alpha_h \in [0, 1]. \quad (4)$$

Two variants are considered: a Global Weighted Blend (GWB) using a single weight across all horizons, and a Local Weighted Blend (LWB) with horizon-specific weights.

The two variants are considered to investigate how model combination should account for horizon-dependent performance differences. The Global Weighted Blend (GWB) represents a simple baseline approach, where a single weight is used for all forecast horizons, assuming that one model combination is sufficient across the entire prediction range. In contrast, the Local Weighted Blend (LWB) allows different weights for each horizon, reflecting the observation that model performance varies depending on the prediction step. By comparing these two variants, we can evaluate whether a single global combination is sufficient or whether adapting the weights to each horizon leads to improved forecasting accuracy.

The GWB applies a single blending coefficient:

$$\hat{y}_{t,h}^{(\text{GWB})} = \alpha \hat{y}_{t,h}^{(A)} + (1 - \alpha) \hat{y}_{t,h}^{(B)}, \quad \alpha \in [0, 1]. \quad (5)$$

The optimal  $\alpha^*$  is selected by minimizing validation RMSE aggregated across all horizons:

$$\alpha^* = \arg \min_{\alpha \in [0,1]} \sqrt{\frac{1}{N \cdot H} \sum_{t=1}^N \sum_{h=1}^H \left( y_{t,h} - \hat{y}_{t,h}^{(\text{GWB})}(\alpha) \right)^2} \quad (6)$$

where  $H=24$ .

The LWB allows horizon-specific blending:

$$\hat{y}_{t,h}^{(\text{LWB})} = \alpha_h \hat{y}_{t,h}^{(A)} + (1 - \alpha_h) \hat{y}_{t,h}^{(B)}, \quad \alpha_h \in [0, 1]. \quad (7)$$

Each  $\alpha_h^*$  is optimized independently on the validation set:

$$\alpha_h^* = \arg \min_{\alpha_h \in [0,1]} \sqrt{\frac{1}{N} \sum_{t=1}^N \left( y_{t,h} - \hat{y}_{t,h}^{(\text{LWB})}(\alpha_h) \right)^2}. \quad (8)$$

Finally, we selected the base models separately for each parking facility to account for differences in capacity and demand dynamics. The models were ranked by RMSE aggregated across all horizons, the top four retained, and the two most frequently best-performing models across horizons selected as the complementary pair  $(A, B)$ . The

selected pairs were Linear Regression+Encoder-Decoder (Seq2Seq) for Lot 38 and Gradient Boosting+Random Forest for Lot 634. Both GWB and LWB variants were evaluated, with the Local Weighted Blend achieving the best overall performance for both facilities.

### 4.3 Evaluation Metrics

To evaluate the model performance, we used standard regression metrics, which are commonly applied in multi-step time series forecasting. Let  $N$  denote the number of test samples and  $H=24$  the prediction horizon. For each sample  $i$  and horizon  $h$ ,  $y_{i,h}$  represents the observed number of free parking spaces and  $\hat{y}_{i,h}$  the corresponding prediction. All metrics are computed over the full set of  $N \times H$  predictions.

Specifically, Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE) are used to quantify average prediction error magnitude, while the coefficient of determination ( $R^2$ ) measures the proportion of variance in the observed data explained by the model. Lower values of MSE, RMSE, and MAE indicate better predictive accuracy, whereas higher  $R^2$  values indicate improved goodness of fit.

## 5 Results

This section reports the experimental results of the proposed forecasting framework for both parking facilities. We begin with a baseline model evaluation using standard error metrics, followed by a horizon-wise analysis of predictive performance. Finally, the proposed hybrid models are assessed against the strongest baselines to demonstrate their effectiveness in capturing horizon-dependent behavior.

### 5.1 Baseline Model Performance

This subsection summarizes the performance of all baseline models prior to hybridization using RMSE, MAE, and  $R^2$  (Section 4.3), with results reported separately for the two parking facilities (Table 2 and 3). The goal is to identify a reduced set of strong baseline models for subsequent horizon-wise analysis.

Baseline performance differs substantially between the two facilities. For Lot 38, deep learning models achieve the best overall accuracy, with the Encoder-Decoder yielding the lowest RMSE and highest  $R^2$ , followed closely by CNN-LSTM. Linear and Ridge Regression also perform competitively, while most tree-based ensemble models exhibit higher errors and negative  $R^2$  values.

In contrast, results for Lot 634 are dominated by classic ensemble methods. Gradient Boosting achieves the

**Table 2** Baseline model performance on Lot 38 before hybridization.

| Model             | RMSE   | MAE    | R <sup>2</sup> |
|-------------------|--------|--------|----------------|
| Encoder-Decoder   | 11.115 | 9.345  | 0.687          |
| CNN-LSTM          | 11.830 | 9.924  | 0.646          |
| Linear Regression | 12.226 | 8.567  | 0.620          |
| Ridge Regression  | 12.344 | 8.659  | 0.613          |
| LSTM              | 12.532 | 10.543 | 0.603          |
| BiLSTM            | 13.583 | 11.388 | 0.533          |
| GRU               | 14.231 | 12.012 | 0.487          |
| CNN-GRU           | 19.119 | 16.115 | 0.075          |
| Gradient Boosting | 20.480 | 16.243 | −0.064         |
| Random Forest     | 20.692 | 16.002 | −0.085         |
| XGBoost           | 20.733 | 16.266 | −0.091         |
| Decision Tree     | 20.983 | 16.548 | −0.117         |
| SVR               | 27.343 | 24.310 | −0.893         |

**Table 3** Baseline model performance on Lot 634 before hybridization.

| Model             | RMSE  | MAE   | R <sup>2</sup> |
|-------------------|-------|-------|----------------|
| Gradient Boosting | 5.217 | 3.374 | 0.914          |
| Random Forest     | 5.236 | 3.432 | 0.914          |
| XGBoost           | 5.290 | 3.413 | 0.912          |
| Ridge Regression  | 5.724 | 3.792 | 0.897          |
| Linear Regression | 5.777 | 3.854 | 0.895          |
| CNN-GRU           | 6.139 | 4.204 | 0.882          |
| Encoder-Decoder   | 6.303 | 4.269 | 0.875          |
| GRU               | 6.447 | 4.549 | 0.870          |
| LSTM              | 6.479 | 4.560 | 0.868          |
| SVR               | 6.500 | 4.511 | 0.867          |
| CNN-LSTM          | 6.655 | 4.691 | 0.861          |
| BiLSTM            | 6.713 | 4.526 | 0.859          |
| Decision Tree     | 7.477 | 4.840 | 0.825          |

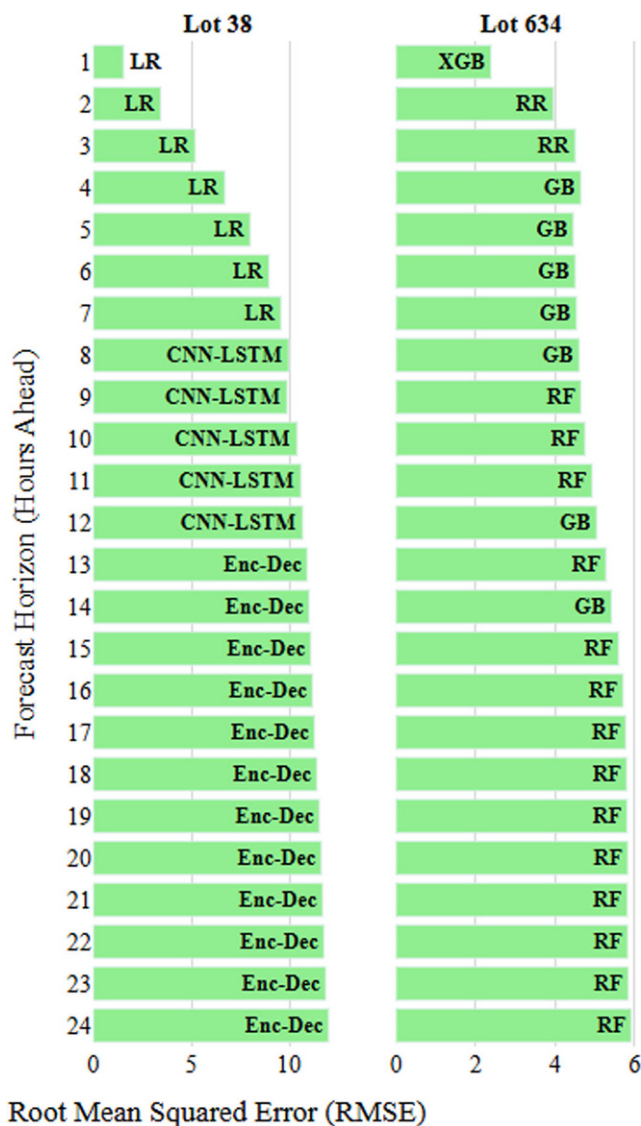
strongest performance, followed closely by Random Forest and XGBoost, while Ridge Regression also ranks among the top-performing models. Deep learning approaches generally perform worse on this lot, with Decision Tree yielding the weakest results.

For each facility, the four best-performing models according to overall RMSE and  $R^2$  are retained and analyzed in a horizon-wise manner, forming the basis for the hybrid model selection strategy.

## 5.2 Horizon-wise Top Models Performance

Figure 2 illustrates the best-performing baseline model at each forecast horizon for Lot 38 and Lot 634, revealing a clear horizon-dependent pattern in predictive accuracy.

For Lot 38, Linear Regression yields the lowest RMSE at short horizons (hours 1–5), CNN-LSTM dominates intermediate horizons (hours 6–12), and the Encoder-Decoder model consistently performs best at longer horizons (hours

**Fig. 2** Comparison of Best Models and RMSE per Forecast Horizon for the top-selected models on Lot 38 and Lot 634

13–24). This progression reflects a shift from local trend exploitation to models capable of capturing longer-term temporal dependencies.

For Lot 634, early horizons are best captured by boosting-based or linear models, while Random Forest achieves the lowest RMSE across most medium and long horizons. Compared to Lot 38, model transitions occur more frequently, indicating more heterogeneous demand dynamics.

Overall, these results confirm that no single baseline model is optimal across all forecast steps, directly motivating the proposed horizon-aware hybrid forecasting strategy.

**Table 4** Comparison of hybrid models against the top four baseline models on Lot 38 (aggregated over the 24-hour horizon).

| Model             | RMSE   | MAE   | $R^2$ |
|-------------------|--------|-------|-------|
| LWBH              | 10.573 | 8.048 | 0.716 |
| GWBH              | 10.783 | 8.282 | 0.704 |
| Encoder-Decoder   | 11.115 | 9.345 | 0.687 |
| CNN-LSTM          | 11.829 | 9.923 | 0.645 |
| Linear Regression | 12.226 | 8.566 | 0.620 |
| Ridge Regression  | 12.344 | 8.659 | 0.613 |

**Table 5** Comparison of hybrid models against the top four baseline models on Lot 634 (aggregated over the 24-hour horizon).

| Model             | RMSE  | MAE   | $R^2$ |
|-------------------|-------|-------|-------|
| LWBH              | 5.060 | 3.265 | 0.919 |
| GWBH              | 5.067 | 3.288 | 0.919 |
| Gradient Boosting | 5.217 | 3.374 | 0.914 |
| Random Forest     | 5.235 | 3.431 | 0.913 |
| XGBoost           | 5.290 | 3.413 | 0.912 |
| Ridge Regression  | 5.724 | 3.791 | 0.897 |

### 5.3 Hybrid Models Performance Comparison

This part compares the proposed hybrid forecasting approaches with the strongest baseline models using aggregate error metrics over the full 24-hour horizon (Table 4 and 5).

Across both parking facilities, the hybrid models consistently outperform all individual baselines. In particular, the Local Weighted Blend Hybrid (LWB) achieves the lowest RMSE and MAE and the highest  $R^2$ , indicating the most accurate overall performance across the forecast horizon. The Global Weighted Blend Hybrid (GWB) also improves upon all baseline models, but remains slightly inferior to the horizon-specific weighting employed by the LWB.

For Lot 38, the LWB yields the best overall results (RMSE=10.573, MAE=8.049,  $R^2$ =0.716), outperforming both linear and deep learning baselines. A similar trend

is observed for Lot 634, where the LWB achieves the strongest performance (RMSE=5.060, MAE=3.265,  $R^2$ =0.920), followed closely by the GWB and ensemble-based models.

In addition to the aggregate evaluation, Fig. 3 and 4 present a horizon-wise comparison of hybrid and baseline models across the 24-hour forecast window. The results reveal a clear horizon-dependent performance pattern and further confirm the advantage of the proposed hybrid approaches.

For Lot 38, Linear Regression achieves the lowest error at very short horizons (hours 1–7), reflecting its strength in capturing local trends. However, from hour 8 onward, the Local Weighted Blend (LWB) consistently outperforms all other models, maintaining the lowest RMSE across medium and long-term horizons. This demonstrates the ability of the hybrid approach to adapt to changing prediction dynamics as the forecast horizon increases.

For Lot 634, the superiority of the LWB is even more pronounced, as it achieves the lowest RMSE across all forecast horizons. Both baseline models (Random Forest and Gradient Boosting) as well as the Global Weighted Blend (GWB) show competitive performance at certain horizons, but remain consistently inferior to the horizon-specific weighting employed by the LWB.

Overall, these findings indicate that while individual models may perform well at specific horizons, the proposed Local Weighted Blend provides a more robust and consistently accurate solution across the entire prediction range. This horizon-wise analysis complements the aggregate results and highlights the effectiveness of incorporating horizon-dependent model behavior into the forecasting process.

## 6 Discussion

This study examined short-term (1–24h) parking availability forecasting using sensor-based occupancy data, with

Horizon-wise RMSE Comparison of Hybrid and Baseline Models – Lot 38

| Model/Horizon     | 1     | 2    | 3    | 4    | 5    | 6     | 7     | 8     | 9     | 10    | 11    | 12    | 13    | 14    | 15    | 16    | 17    | 18    | 19    | 20    | 21    | 22    | 23    | 24    |
|-------------------|-------|------|------|------|------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| <b>LWB</b>        | 1.54  | 4.34 | 6.32 | 7.93 | 8.97 | 9.64  | 10.27 | 10.84 | 11.44 | 11.74 | 11.86 | 11.93 | 12.04 | 12.07 | 12.09 | 11.91 | 12.01 | 11.85 | 11.98 | 11.10 | 11.23 | 11.96 | 12.03 | 12.37 |
| <b>GWB</b>        | 1.71  | 4.40 | 6.60 | 8.34 | 9.43 | 9.82  | 10.42 | 11.01 | 11.58 | 11.81 | 11.95 | 12.06 | 12.16 | 12.19 | 12.20 | 12.07 | 12.03 | 11.90 | 11.27 | 11.29 | 11.77 | 12.07 | 12.11 | 12.45 |
| <b>LR</b>         | 1.54  | 3.44 | 5.20 | 6.71 | 8.02 | 9.18  | 10.20 | 11.11 | 11.92 | 12.60 | 13.14 | 13.53 | 13.79 | 13.97 | 14.13 | 14.25 | 14.37 | 14.43 | 14.46 | 14.47 | 14.48 | 14.52 | 14.58 | 14.64 |
| <b>ED</b>         | 10.87 | 9.26 | 9.04 | 9.39 | 9.86 | 10.33 | 10.78 | 11.19 | 11.54 | 11.82 | 12.03 | 12.17 | 12.24 | 12.26 | 12.21 | 12.15 | 12.08 | 12.04 | 12.04 | 12.07 | 12.14 | 12.24 | 12.36 | 12.50 |
| <b>Best Model</b> | LR    | LR   | LR   | LR   | LR   | LR    | LR    | LWB   | LWB   | LWB   | LWB   | LWB   | LWB   | LWB   | LWB   | LWB   | LWB   | LWB   | LWB   | LWB   | LWB   | LWB   | LWB   | LWB   |

**Fig. 3** Horizon-wise RMSE comparison of hybrid (LWB, GWB) and baseline (Linear Regression, Encoder-Decoder) models across the 24-hour forecast horizon for Lot 38. The Local Weighted Blend (LWB) achieves the lowest error for most horizons, while Linear Regression (LR) performs best at very short horizons

Horizon-wise RMSE Comparison of Hybrid and Baseline Models – Lot 634

| Model/Horizon     | 1    | 2    | 3    | 4    | 5    | 6    | 7    | 8    | 9    | 10   | 11   | 12   | 13   | 14   | 15   | 16   | 17   | 18   | 19   | 20   | 21   | 22   | 23   | 24   |
|-------------------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|
| <b>LWB</b>        | 2.41 | 4.22 | 4.69 | 4.50 | 4.40 | 4.34 | 4.39 | 4.43 | 4.53 | 4.67 | 4.81 | 4.98 | 5.19 | 5.33 | 5.53 | 5.63 | 5.68 | 5.73 | 5.74 | 5.76 | 5.74 | 5.77 | 5.75 | 5.76 |
| <b>GWB</b>        | 2.41 | 4.24 | 4.70 | 4.60 | 4.43 | 4.40 | 4.42 | 4.51 | 4.58 | 4.70 | 4.85 | 5.02 | 5.26 | 5.39 | 5.56 | 5.68 | 5.70 | 5.76 | 5.77 | 5.79 | 5.77 | 5.80 | 5.79 | 5.80 |
| <b>RF</b>         | 3.86 | 4.73 | 4.98 | 4.75 | 4.63 | 4.58 | 4.61 | 4.64 | 4.65 | 4.75 | 4.93 | 5.11 | 5.28 | 5.47 | 5.60 | 5.71 | 5.78 | 5.80 | 5.81 | 5.83 | 5.82 | 5.83 | 5.84 | 5.92 |
| <b>GB</b>         | 2.42 | 4.26 | 4.77 | 4.65 | 4.46 | 4.51 | 4.54 | 4.61 | 4.80 | 4.94 | 4.95 | 5.05 | 5.33 | 5.42 | 5.73 | 5.83 | 5.89 | 5.90 | 5.86 | 5.85 | 5.90 | 6.02 | 5.98 | 5.96 |
| <b>Best Model</b> | LWB  | LWB  | LWB  | LWB  | LWB  | LWB  | LWB  | LWB  | LWB  | LWB  | LWB  | LWB  | LWB  | LWB  | LWB  | LWB  | LWB  | LWB  | LWB  | LWB  | LWB  | LWB  | LWB  | LWB  |

**Fig. 4** Horizon-wise RMSE comparison of hybrid (LWB, GWB) and baseline (Random Forest, Gradient Boosting) models across the 24-hour forecast horizon for Lot 634. The Local Weighted Blend (LWB) consistently achieves the lowest error across all horizons, outperforming both baseline models and the Global Weighted Blend (GWB)

a focus on horizon-dependent predictive accuracy. Results show that no single model achieves uniformly low error across all forecast steps and that model performance varies substantially with prediction horizon, which has direct implications for model selection in practical parking information systems.

Across both parking facilities, baseline results reveal complementary strengths among model families. For Lot 38, simpler models perform best at short horizons, CNN-LSTM models dominate at medium horizons, and the Encoder-Decoder architecture achieves the lowest error at longer horizons, reflecting a transition from local extrapolation to long-term sequence modeling. A similar horizon-dependent pattern is observed for Lot 634, where classic ensemble methods dominate overall, but different models are optimal at different lead times. These findings confirm prior observations that model complexity alone does not guarantee superior performance and that the optimal predictor depends on forecast horizon.

The proposed hybrid forecasting approach directly addresses this limitation by combining complementary models. While the Global Weighted Blend improves robustness over individual baselines, the Local Weighted Blend consistently achieves the best performance across both parking facilities. The improvements observed across RMSE, MAE, and  $R^2$  demonstrate that horizon-specific weighting effectively exploits model strengths where they are most relevant.

Qualitative inspection of representative 24-hour forecast trajectories further shows that hybrid models track observed availability more closely than individual baselines, particularly at longer horizons where single models tend to exhibit systematic deviations. This is especially important for decision-support applications that require reliable next-day availability trajectories rather than accuracy at isolated time steps.

Several limitations remain: The evaluation is limited to two parking facilities within a single city, which most of

the time have at least some availability. Our approach might yield different resulting models for other parking lots. Additional contextual factors such as special events or traffic disruptions were not explicitly modeled and could further influence our results. Moreover, the hybrid framework currently combines only two base predictors. Future work should extend the approach to multi-model ensembles, incorporate richer contextual data, and evaluate transferability across cities and parking types, as well as assess the impact of horizon-aware forecasts on real-world driver behavior.

## 7 Conclusion and Future Work

This work presents a horizon-aware hybrid forecasting approach for smart parking systems using sensor-based occupancy data. By systematically evaluating a diverse set of classic machine learning and deep learning models under a unified experimental protocol in two real-world scenarios, the study demonstrated that predictive accuracy varies substantially across forecast horizons and that no single model is universally optimal over a 24-hour range. To address this challenge, we introduced a Global Weighted Blend Hybrid and a Local Weighted Blend Hybrid forecasting strategy. Experimental results on two parking facilities showed that both hybrid variants outperform individual baseline models, with the Local Weighted Blend Hybrid achieving the best overall performance. These results suggest that explicitly accounting for horizon-dependent model behavior could lead to more accurate and robust multi-step forecasts.

Beyond methodological contributions, the proposed approach was developed and evaluated within an end-to-end data engineering pipeline, emphasizing deployment feasibility and decision-support relevance. By translating multi-horizon predictions into interpretable next-day availability trajectories, the system could support informed arrival-time planning and enhance the practical usefulness of parking information services.

To further study the feasibility of this approach, we plan to integrate this approach into an adaptive architecture for real-time inference. This architecture would manage multiple machine learning models and dynamically adapt the data pipelines to the workload of downstream tasks (here: different prediction horizons). Since using context information could provide higher accuracy than any single-model approach, we propose to structure the system as an adaptive system implementing the MAPE-K loop (Monitor, Analyze, Plan, Execute, over a shared Knowledge base), a reference model for engineering feedback control in autonomic computing [36].

Overall, this study highlights the importance of moving beyond single-model forecasting in multi-step prediction settings and demonstrates that horizon-aware hybridization could be a simple yet effective strategy for improving predictive performance in real-world smart city applications.

**Funding** Open Access funding enabled and organized by Projekt DEAL.

**Data availability** All data supporting the findings of this work are included in the article.

**Open Access** Dieser Artikel wird unter der Creative Commons Namensnennung 4.0 International Lizenz veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden. Die in diesem Artikel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen. Weitere Details zur Lizenz entnehmen Sie bitte der Lizenzinformation auf <http://creativecommons.org/licenses/by/4.0/deed.de>.

## References

- Hall P (2014) Cities of tomorrow: An intellectual history of urban planning and design since 1880. John Wiley & Sons
- D N, L A, D D et al (2026) Challenges for Federated Crowd Management in Smart Cities. Companion Proceedings of the 27th International Conference on Distributed Computing and Networking. ACM Other conferences, Association for Computing Machinery. <https://doi.org/10.1145/3737611.3776621>
- Dustdar S, Nastić S, Šćekić O (2017) Smart cities. The Internet of Things, People and Systems <https://doi.org/10.1007/978-3-319-60030-7>
- Lu J, Liu A, Dong F et al (2018) Learning under concept drift: A review. IEEE Trans Knowl Data Eng 31(12):2346–2363. <https://doi.org/10.1109/TKDE.2018.2876857>
- Ozeki R, Yonekura H, Rizk H et al (2024) Privacy Preserved Taxi Demand Prediction System for Distributed Data. Proceedings of the 32nd ACM International Conference on Advances in Geographic Information Systems. <https://doi.org/10.1145/3678717.3691234>
- Ma C, Huang X, Li J (2024) A review of research on urban parking prediction. Journal of Traffic and Transportation Engineering 11(4):700–720. <https://doi.org/10.1016/j.jtte.2023.11.004>
- Polycarpou E, Lambrinos L, Protopapadakis E (2013) Smart parking solutions for urban areas. 2013 IEEE 14th International Symposium on “A World of Wireless, Mobile and Multimedia Networks” (WoWMoM). <https://doi.org/10.1109/WoWMoM.2013.6583499>
- Bischoff J, Nagel K (2017) Integrating explicit parking search into a transport simulation. Procedia Comput Sci 109:881–886. <https://doi.org/10.1016/j.procs.2017.05.414>
- Caicedo F (2010) Real-time parking information management to reduce search time, vehicle displacement and emissions. Transp Res Part F Traffic Psychol Behav 15(4):228–234. <https://doi.org/10.1016/j.trd.2010.02.008>
- Lyu M, Ji Y, Kuai C et al (2024) Short-term prediction of on-street parking occupancy using multivariate variable based on deep learning. Journal of Traffic and Transportation Engineering 11(1):28–40. <https://doi.org/10.1016/j.jtte.2022.05.004>
- Awan FM, Saleem Y, Minerva R et al (2020) A comparative analysis of machine/deep learning models for parking space availability prediction. Sensors 20(1):322. <https://doi.org/10.3390/s20010322>
- Channamallu SS, Kermanshachi S, Rosenberger JM et al (2023) Parking occupancy prediction and analysis - a comprehensive study. Transportation Research Procedia 73:297–304. <https://doi.org/10.1016/j.trpro.2023.11.921>
- Bilotta S, Ipsaro Palesi LA, Nesi P (2023) Predicting free parking slots via deep learning in short-mid terms explaining temporal impact of features. IEEE Access 11:101678–101693. <https://doi.org/10.1109/ACCESS.2023.3314660>
- Zhao Z, Zhang Y, Zhang Y (2020) A comparative study of parking occupancy prediction methods considering parking type and parking scale. Journal of Advanced Transportation 2020:1–12. <https://doi.org/10.1155/2020/5624586>
- Elomiya A, Křupka J, Jovčić S et al (2024) Enhanced prediction of parking occupancy through fusion of adaptive neuro-fuzzy inference system and deep learning models. Eng Appl Artif Intell 129:107670. <https://doi.org/10.1016/j.engappai.2023.107670>
- Yang S, Ma W, Pi X, et al. (2019) A deep learning approach to real-time parking occupancy prediction in spatio-temporal networks incorporating multiple spatio-temporal data sources. <https://doi.org/10.48550/ARXIV.1901.06758>, version Number: 5
- Vlahogianni EI, Kepaptsoglou K, Tsetos V et al (2016) A real-time parking prediction system for smart cities. Journal of Intelligent Transportation Systems 20(2):192–204. <https://doi.org/10.1080/15472450.2015.1037955>
- Domenech F, Carot V, Portales C et al (2025) AI-based prediction models for urban parking availability: A case study of valencia. SAUC 11(7):1–19. <https://doi.org/10.62161/sauc.v11.5990>
- Channamallu S S, Kermanshachi S, J MR et al (2024) Enhancing urban parking efficiency through machine learning model integration. IEEE Access 12:81338–81347. <https://doi.org/10.1109/ACCESS.2024.3411391>
- Koumetis Tekouabou SC, Abdellaoui Alaoui EA, Cherif W et al (2022) Improving parking availability prediction in smart cities with IoT and ensemble-based model. J King Saud Univ Comput Inf Sci 34(3):687–697. <https://doi.org/10.1016/j.jksuci.2020.01.008>
- Channamallu SS, Padavala VK, Kermanshachi S et al (2023) Examining parking occupancy prediction models: a comparative analysis. Transportation Research Procedia 73:281–288. <https://doi.org/10.1016/j.trpro.2023.11.919>
- Vakrinou K, Mantouka EG, Vlahogianni EI (2025) Leveraging multisource data for off-street parking occupancy prediction: A twofold analysis. International Journal of Transportation Science and Technology S204604302500036X. <https://doi.org/10.1016/j.ijtst.2025.02.005>

23. Hastie T, Tibshirani R, Friedman J et al (2009) The elements of statistical learning
24. Chen T, Guestrin C (2016) XGBoost: A scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. <https://doi.org/10.1145/2939672.2939785>
25. Han C, Park H, Kim Y et al (2022) Hybrid cnn-lstm based time series data prediction model study. IEEE/ACIS International Conference on Big Data, Cloud Computing, and Data Science Engineering. [https://doi.org/10.1007/978-3-031-19608-9\\_4](https://doi.org/10.1007/978-3-031-19608-9_4)
26. Hoerl AE, Kennard RW (1970) Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics* 12(1):55–67. <https://doi.org/10.1080/00401706.1970.10488634>
27. Loh W (2011) Classification and regression trees. *WIREs Data Min & Knowl* 1(1):14–23. <https://doi.org/10.1002/widm.8>
28. Breiman L (2001) Random forests. *Machine Learning* 45(1):5–32. <https://doi.org/10.1023/A:1010933404324>
29. Friedman J (2001) Greedy function approximation: A gradient boosting machine. *The Annals of Statistics* 29(5):1189–1232
30. Smola AJ, Schölkopf B (2004) A tutorial on support vector regression. *Statistics and Computing* 14(3):199–222. <https://doi.org/10.1023/B:STCO.0000035301.49549.88>
31. Hochreiter S, Schmidhuber J (1997) Long short-term memory. *Neural Computation* 9(8):1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
32. Cho K, van Merriënboer B, Gulcehre C, et al. (2014) Learning phrase representations using RNN encoder-decoder for statistical machine translation. <https://doi.org/10.48550/ARXIV.1406.1078>, version Number: 3
33. Schuster M, Paliwal K (1997) Bidirectional recurrent neural networks. *IEEE Trans Signal Process* 45(11):2673–2681. <https://doi.org/10.1109/78.650093>
34. Sutskever I, Vinyals O, Le QV (2014) Sequence to sequence learning with neural networks. <https://doi.org/10.48550/ARXIV.1409.3215>, version Number: 3
35. Zhang F, Shang K, Yan L et al (2024) Prediction of parking space availability using improved MAT-LSTM network. *ISPRS Int J Geoinf* 13(5):151. <https://doi.org/10.3390/ijgi13050151>
36. Casimiro M, Romano P, Garlan D, et al. (2021) Self-Adaptation for Machine Learning Based Systems. In: Heinrich R, Mirandola R, Weyns D (eds) ECSCA 2021 Companion Volume, Virtual (originally: Växjö, Sweden), 13–17 September, 2021, CEUR Workshop Proceedings, vol 2978. CEUR-WS.org. <https://ceur-ws.org/Vol-2978/saml-paper6.pdf>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.