Short and Long Term Bike Rental Forecasting

Oskar Kocjančič[*]
oskar.kocjancic@gmail.com
Jožef Stefan Institute
Ljubljana, Slovenia

Martin Žnidaršič
martin.znidarsic@ijs.si
Jožef Stefan Institute
Ljubljana, Slovenia

Abstract

This paper describes the challenges and outcomes of forecasting bike rentals in a Slovenian urban bike-sharing system, focusing on the impact of data sparsity and the inclusion of external variables. We address two distinct forecasting tasks: short-horizon, one-day-ahead predictions for individual rental stations, and long-horizon, 90-day forecasts for the total rental volume. Various machine learning models were employed and evaluated in this context. We also analyzed the trade-off between using longer historical data versus shorter, weather-enriched data to improve predictive accuracy. The findings indicate a clear correlation between data sparsity at the station level and predictive performance. While the inclusion of weather data provides a modest improvement for both short-horizon and long-horizon forecasts, the overall quality of the sparse and noisy data appears to limit the potential gains from more complex modeling approaches.

Keywords

bike-sharing, forecasting, time series, data sparsity, machine learning, deep learning, weather data

1 Introduction

Predicting rental patterns of urban bike-sharing systems is challenging due to complex dynamics, including strong seasonality and trends, as well as dependence on external variables such as weather and calendar effects. Furthermore, data sparsity, particularly at the individual station level, presents a significant obstacle to building reliable predictive models. By accurately predicting bike demand, operators can improve redistribution and station availability, fostering a more reliable and sustainable urban mobility system.

This paper addresses these challenges by investigating two distinct forecasting tasks using a real-world dataset from a Slovenian city. First, we examine short-horizon, one-day-ahead predictions for individual stations to quantify the impact of data sparsity on forecastability. Second, we evaluate the accuracy of 90-day long-horizon forecasts for the total rental volume aggregated across all stations. We compare a suite of models, including classical machine learning approaches and LSTM neural networks [5], and explicitly analyze the trade-off between using longer historical data versus shorter, weather-enriched data to improve predictive accuracy. This work aims to help the bike-sharing systems to improve operational efficiency, reduce bike shortages, and inform city planning initiatives related to sustainable transportation.

[*]Both authors contributed equally to this research.

Information Society 2025, Ljubljana, Slovenia

https://doi.org/10.70314/is.2025.sikdd.7

Prior studies on bicycle rental forecasting often use the Washington, D.C. dataset [4]. Du et al. [2] addressed long-horizon prediction, while Karunanithi et al. [6] focused on short-horizon forecasting, both achieving results comparable to ours. In contrast, our dataset differs substantially by including station-level information, which enables per-station forecasting. We tackle both short- and long-horizon tasks, as well as the analysis of the impact of exogenous weather variables.

2 Data

The dataset we used originates from a public bicycle rental service in a Slovenian city. It contains daily rental counts for individual stations within the municipality, covering the period from January 1, 2021, to May 15, 2025. Although the dataset also records bike return counts, our work focuses exclusively on rentals.

2.1 Features

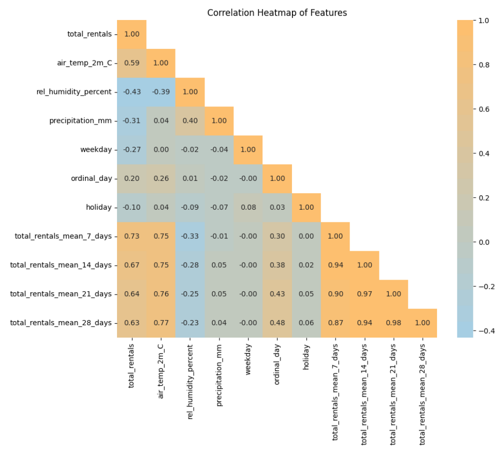

Figure 1: Pearson correlation coefficients of our features

Dependent Variable: The target feature we are forecasting.

- **total_rentals:** The total daily number of bike rentals. Based on the task, this is either the total count across all stations or per-station bike rental count.

Independent Variables: The features used for prediction.

- **Temporal Features:**
  - **date:** The specific date.
  - **ordinal_day:** The day number within the year.
  - **weekday:** A category for the day of the week.
  - **holiday:** Indicator (0 or 1) if the day is a holiday.
- **Weather-Related Features:** *Note: Our weather data only spans the date range of 2024-01-01 to 2025-05-14*
  - **air_temp_2m_C:** Air temperature.
  - **rel_humidity_percent:** The relative humidity.
  - **precipitation_mm:** The precipitation per square meter.

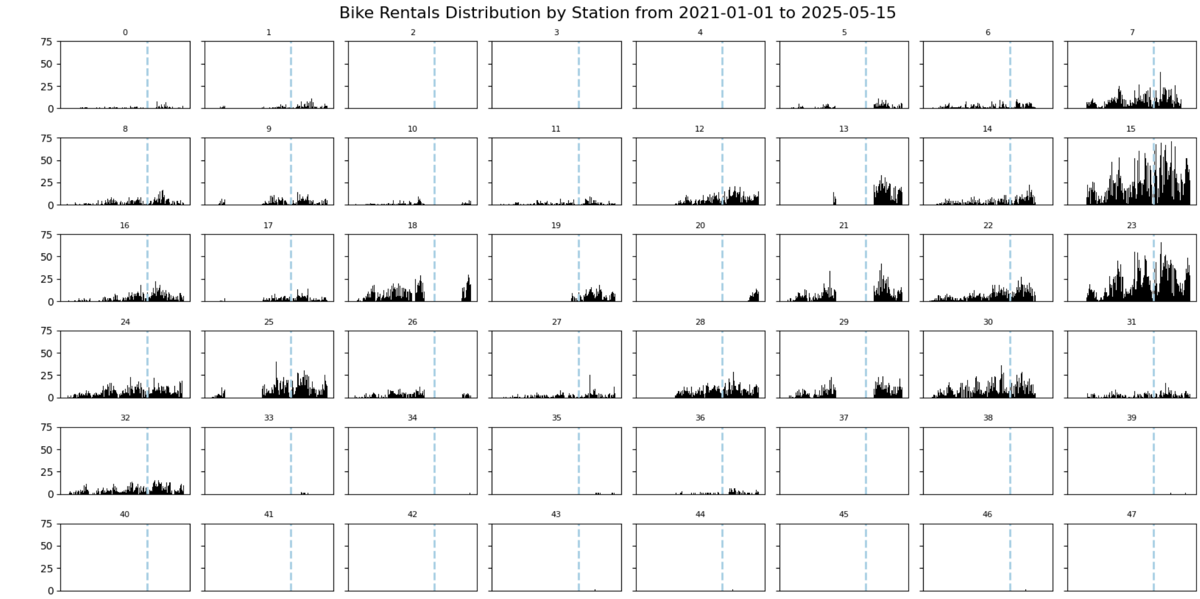

Figure 2: Distribution of bike rentals across all stations. The vertical blue line indicates the start of the year 2024.

2.2 Data Preprocessing

The dataset structure prevented distinguishing missing values from true zeros (i.e., days when no rentals occurred), so all empty or null entries were treated as zeros. This resulted in sparsity for some stations, in which many entries had little information on rental activity. To prevent this impacting our analysis, we excluded those with more than 33% zero entries, retaining 25 stations out of the original 48. For the machine learning methods described later, we also implemented a set of **lagged features:**

- **total_rentals_mean_7_days:** Average rental count over the 7 days preceding the current data point.
- **total_rentals_mean_14_days:** Average rental count over the 14 days preceding the current data point.
- **total_rentals_mean_21_days:** Average rental count over the 21 days preceding the current data point.
- **total_rentals_mean_28_days:** Average rental count over the 28 days preceding the current data point.

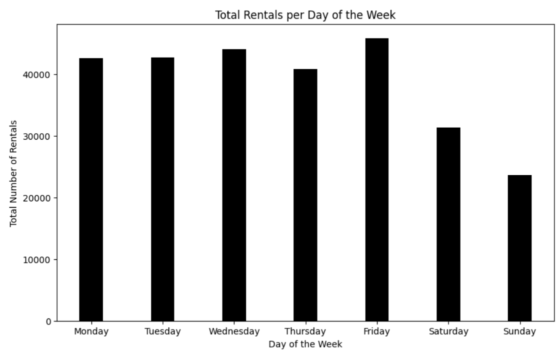

Figure 3: Rentals per day of the week

2.3 Exploratory Data Analysis

The data exhibits pronounced weekly and monthly seasonalities, as well as non-stationarity, as illustrated in Figures 3 and 4. Annual patterns show rental activity declining in winter, rising in spring, peaking in summer, and gradually decreasing in autumn, with weekends consistently exhibiting lower rental counts. Anomalous behavior was observed in the winter of 2024, when rental counts were markedly higher than typical seasonal levels.

The Pearson correlation coefficients (Figure 1) between features related to bicycle rentals indicate that the number of daily rentals (`total_rentals`) is strongly and positively associated with recent rental trends, as reflected by correlations of 0.73, 0.67, 0.64, and 0.63 with the 7-, 14-, 21-, and 28-day moving averages, respectively. A strong positive correlation is also observed with air temperature (0.59), whereas moderate negative correlations are found with relative humidity (-0.43) and precipitation (-0.31), suggesting that rentals are more frequent on warm, dry days. Weaker associations are present with the day of the week (-0.27) and holiday status (-0.10). As expected, the moving average features exhibit high intercorrelation (e.g., 0.94 between the 7- and 14-day means) due to their overlapping calculation windows.

3 Experiments

This study pursued two primary objectives. First, we examined the feasibility of forecasting bicycle rentals one day in advance and evaluated how forecastability varies across stations with different data sparsity. Second, we investigated long-horizon forecasting over a 90-day period, focusing exclusively on predicting the total number of rentals. In this task, standard machine learning models were trained on historical data and then used recursively to generate forecasts for the entire period. Due to this setup, the results for **DS_W** suffer from data leakage. Specifically, a single model is trained using past rental counts and future weather information, so, for example, predicting rentals in July involves access to the actual recorded weather conditions for that month, which artificially improves performance.

3.1 Training and Test Data Split

Because the available weather data was limited to the years 2024 and 2025, while the rental dataset spanned from 2021 onward,

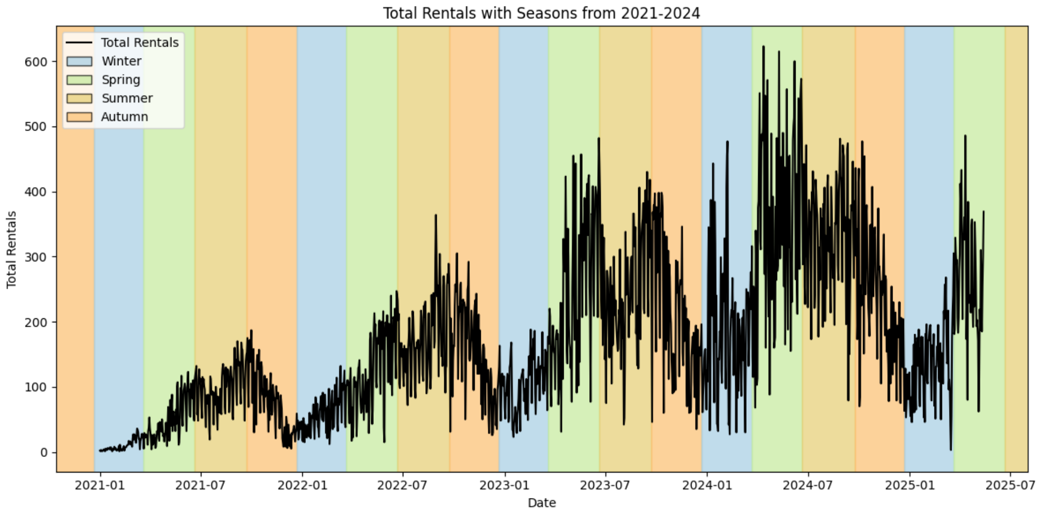

Figure 4: Bike rental data with temperate seasons

we constructed three distinct datasets. Here, each *entry* corresponds to a single day and includes rental data for all stations. The first dataset, **DS_W**, combined rental and weather data (498 entries). The second, **DS_NO_W**, included only rental data for the same period (498 entries). The third, **DS_FULL**, comprised the complete rental dataset without weather data (1,593 entries).

The data splitting strategy differed in the two tasks. For the station-level one-day-ahead forecasting task, each dataset was divided into 25 subsets, corresponding to individual stations. Within each subset, random sampling was used to split the data into training and testing sets with an 80:20 ratio. The target variable in each subset is the specific station's rental count.

For the long-horizon task, no station-level subdivision was performed, as only total rental counts were modeled. The final 90 days were used as the test set—roughly corresponding to a temperate season—allowing us to assess whether the models capture seasonal patterns in a new period while maintaining realistic temporal separation between training and testing data.

3.2 Models and Algorithms Used

For the long-horizon forecasting task, the **AutoARIMA** model served as the baseline, while for the one-day-ahead forecasting task, the baseline was the **Mean Regressor**, which predicts using the 7-day lag mean.

We evaluated several machine learning models, including **Random Forest** (500 trees, max_features=0.9), **Gradient Boosting** (500 estimators), **Linear Regression**, and **SVM** ($C = 10$, degree=2, $\gamma = 0.1$, linear kernel). The hyperparameters for the Random Forest and SVM models were selected using a grid search optimization procedure; the rest of the models used default parameters. For the Random Forest model, only the `max_features` parameter was tuned.

We additionally tested deep learning approaches: **LSTM** (input size = 96, RMSE loss, 10,000 epochs) and **N-BEATSx** (input size = 96, RMSE loss, 500 epochs).

Training was performed on a laptop equipped with an RTX 3050 GPU (4 GB VRAM), which constrained the range of hyperparameter configurations that could be explored, particularly for the neural network-based approaches.

3.3 Performance evaluation

Model performance was assessed using Root Mean Squared Error (RMSE) and Mean Absolute Percentage Error (MAPE). Additionally, the Relative Root Mean Squared Error (RRMSE)[1] was used to enable inter-station performance comparisons in the one-day-ahead forecasting task. RRMSE is defined as follows:

$$\mathrm{RRMSE} = \frac{\mathrm{RMSE}}{\overline{y}} \quad (1)$$

where $\overline{y}$ is the mean of the target values.

3.4 Results

The results for the one-day-ahead task are presented in Table 1, with station forecastability visualized in Figure 5. The long-horizon task outcomes are presented in Table 2.

4 Discussion and conclusion

For the one-day-ahead forecasting task, a clear correlation exists between station data sparsity (Figure 2) and forecastability (Table 1). Stations with fewer rentals or gaps in data are easier to predict accurately. Interestingly, using the **DS_FULL** dataset—which includes data prior to 2024—can reduce modeling accuracy for certain stations. Including weather features in **DS_W** leads to little or no improvement compared to **DS_NO_W**. For the long-horizon task, including weather data proves beneficial, as both classical machine learning models and neural networks show improved performance (Table 2). However, as described in the Experiments section, the machine learning results on **DS_W** are overly optimistic due to data leakage: the models are trained on historical rental counts while also accessing future weather information during recursive forecasting (e.g., predicting rentals in July uses the actual recorded weather for that month). This is reflected in the comparison with **DS_NO_W**, where classical machine learning methods achieve a 33% mean reduction in MAPE, while neural network approaches show only a 17% mean decrease, suggesting that the apparent benefit of weather data is amplified for classical methods because of this setup. Our results echo [3] where Gradient Boosting models matched or outperformed neural networks on several datasets, demonstrating the effectiveness of simpler models. While neural networks

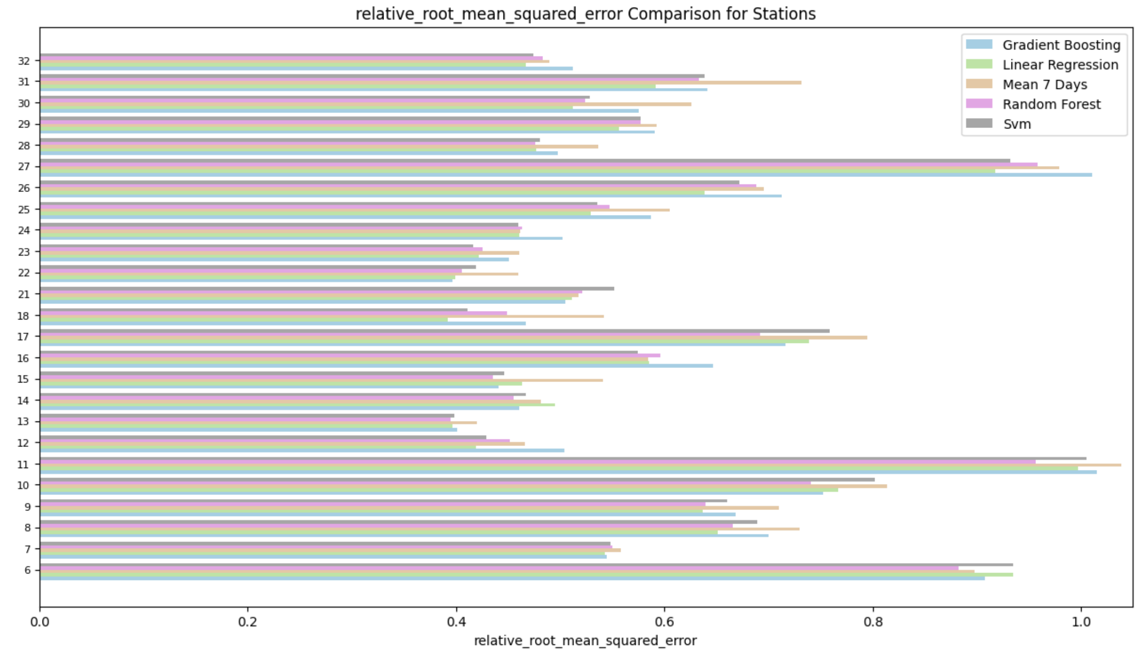

Figure 5: Model performance of one-day-head forecasting for different stations for DS_W

Table 1: Average RRMSE of all models of one-day-ahead forecasting across datasets (RRMSE) and stations.

Station	DS_FULL	DS_NO_W	DS_W
6	0.9210	**0.9097**	0.9116
7	0.5849	**0.5439**	0.5488
8	0.7948	**0.6821**	0.6872
9	**0.6532**	0.6646	0.6631
10	0.9550	**0.7747**	0.7753
11	1.0110	1.0034	**1.0027**
12	0.6028	0.4649	**0.4540**
13	0.6601	**0.4000**	0.4022
14	0.6902	0.4840	**0.4720**
15	0.5218	0.4780	**0.4652**
16	0.7185	0.5984	**0.5975**
17	0.8336	**0.7337**	0.7402
18	0.5274	0.4670	**0.4522**
21	0.5476	0.5218	**0.5215**
22	0.5198	0.4171	**0.4160**
23	0.4783	0.4363	**0.4349**
24	0.4896	0.4760	**0.4696**
25	0.6834	**0.5570**	0.5608
26	**0.6506**	0.6897	0.6812
27	**0.9463**	0.9898	0.9595
28	0.5580	**0.4898**	0.4936
29	0.6008	**0.5761**	0.5788
30	0.5941	**0.5496**	0.5531
31	0.8952	**0.6452**	0.6474
32	0.5453	0.4873	**0.4851**
Average	**0.6793**	**0.6016**	**0.5989**

could potentially benefit from hyperparameter optimization, the same applies to other methods as well. A detailed comparison of different approaches was beyond the scope of this preliminary study but could be explored in future work.

Table 2: Model performance of 90-day forecasting across datasets (RMSE / MAPE)

Model	DS_FULL	DS_NO_W	DS_W
AutoARIMA	120.09 / 0.9525	118.50 / 0.9954	118.50 / 0.9954
Random Forest	108.29 / 0.7153	100.94 / **0.7431**	76.36 / 0.7014
Gradient Boosting	95.17 / 0.7451	94.96 / 0.9584	74.69 / **0.5513**
Linear Regression	**90.29** / 0.9372	84.78 / 1.0816	71.71 / 0.8872
SVR	94.86 / 0.8893	**87.12** / 0.9507	**67.95** / 0.8036
LSTM	112.05 / **0.7133**	125.13 / 0.8494	130.00 / 0.8070
NBEATSx	106.49 / 1.0329	128.90 / 0.9972	117.45 / 0.7246
Average	**103.89 / 0.8551**	**105.76 / 0.9394**	**93.81 / 0.7815**

Acknowledgements

This work was supported in part by the Slovenian Research Agency through core funding for the programme Knowledge Technologies (No. P2-0103) and by the project *KReATIVE*, funded through NetZeroCities under the European Union's Grant Agreement No. HORIZON-RIA-SGA-NZC 101121530. We also thank Tea Tušar for her suggestions regarding data visualization.

References

[1] Shikun Chen and Nguyen Manh Luc. 2022. Rrmse voting regressor: a weighting function based improvement to ensemble regression. *arXiv preprint arXiv:2207.04837*.

[2] Jimmy Du, Rolland He, and Zhivko Zhechev. 2014. Forecasting bike rental demand. *Gebhard, K., & Noland*.

[3] Shereen Elsayed, Daniela Thyssens, Ahmed Rashed, Lars Schmidt-Thieme, and Hadi Samer Jomaa. 2021. Do we really need deep learning models for time series forecasting? *CoRR*, abs/2101.02118. https://arxiv.org/abs/2101.02118 arXiv: 2101.02118.

[4] Hadi Fanaee-Tork. 2012. Bike sharing dataset. Dataset. (2012). https://www.kaggle.com/datasets/marklvl/bike-sharing-dataset.

[5] Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long short-term memory. *Neural computation*, 9, 8, 1735–1780.

[6] Meerah Karunanithi, Parin Chatasawapreeda, and Talha Ali Khan. 2024. A predictive analytics approach for forecasting bike rental demand. *Decision Analytics Journal*, 11, 100482. DOI: https://doi.org/10.1016/j.dajour.2024.100482.