Proceedings of CONF-MLA 2026 Symposium: Learning and Decision Making in Multi Agent Software Systems
DOI: 10.54254/2755-2721/2026.BA36242

Short-Term Bike-Sharing Demand Prediction Using Environmental and Temporal Feature Fusion

Lujing Wang

School of Low-Altitude Technology and Electronic Information, Guangdong University of Science and Technology, Dongguan, China

lujingwang37@gmail.com

Abstract. Accurate hourly bike-sharing demand prediction is important for bicycle rebalancing and resource allocation. This study evaluates the individual and combined contributions of environmental and temporal features to hourly bike-sharing demand prediction using 17,414 hourly records from the London Bike Sharing Dataset. A time-ordered 80:20 split is adopted to mimic forecasting future demand from historical observations. A distance-weighted k-nearest neighbors (KNN) regressor is adopted as the baseline, and a tuned random forest (RF) regressor serves as the main model. RF achieves MAE=173.08 rentals/h, RMSE=304.76 rentals/h, and R^2=0.927, reducing MAE and RMSE by 65.7% and 60.2% compared with KNN. Feature-combination experiments show that environmental features alone have limited explanatory power (R^2=0.132), whereas temporal features capture the main demand structure (R^2=0.868). After feature fusion, MAE and RMSE decrease by 29.0% and 25.6% relative to the temporal-only setting. Visual analyses and permutation importance indicate that temporal variables establish the main demand baseline, while environmental variables provide complementary corrections. The results offer interpretable evidence for feature selection and short-term bike-sharing operations.

Keywords: Bike sharing, demand prediction, feature fusion, random forest, temporal features

1. Introduction

Bike-sharing systems have become an important complement to urban public transport. Their operational efficiency depends heavily on short-term demand prediction because underestimation may cause bicycle shortages in high-demand areas, whereas overestimation may lead to idle inventory and unnecessary rebalancing costs. Hourly demand is affected jointly by commuting rhythms, calendar effects, and weather conditions. Therefore, a reliable forecasting model should capture not only the periodic baseline of travel behavior but also external factors that shift demand away from that baseline.

Previous studies have confirmed the relevance of temporal, meteorological, and spatial factors in bike-sharing demand. El-Assi et al. used station-level bike-sharing data from Toronto and showed that weather and land-use conditions are significantly associated with ridership [1]. Zhou analyzed large-scale bike-sharing records in Chicago and identified clear daily and weekly spatiotemporal

DOI: 10.54254/2755-2721/2026.BA36242

regularities, indicating the importance of temporal structure [2]. Jiao et al. compared several machine-learning models for short-term bike-sharing demand prediction and reported that time, weather, and location-related variables are important predictors [3]. Xie et al. further discussed the relationship between weather conditions and urban bike-sharing trips, while Hu et al. examined demand prediction around urban rail stations, showing the practical value of bike-sharing forecasting for first/last-mile travel connections [4, 5].

Recent studies have also adopted advanced learning architectures to capture complex demand patterns. Li et al. proposed an irregular convolutional neural network for short-term bike-sharing demand forecasting and showed that modeling non-Euclidean regional dependencies improves predictive performance [6]. Rennie et al. analyzed bike-sharing demand outliers and emphasized that unusual weather and transport disturbances can affect forecast reliability [7]. Rochas et al. integrated time, weather, and traffic context into a graph-neural-network framework and reported that adding weather information reduced prediction error by more than 20% under degraded weather conditions [8]. Behroozi and Edrisi also showed that graph convolutional networks can improve station-level bike-sharing travel-demand prediction [9].

Although these studies provide valuable evidence, most existing works treat temporal and environmental variables as a unified input and focus mainly on accuracy improvement, with limited systematic analysis of how these two feature groups separately and jointly shape prediction performance. In addition, highly complex neural networks may achieve high accuracy but often reduce interpretability for practical feature selection and operational decision-making. This study therefore focuses on a controlled feature-combination design rather than proposing a more complex architecture.

This paper investigates hourly bike-sharing demand prediction using environmental-temporal feature fusion. First, three input schemes are constructed: environmental features only, temporal features only, and fused environmental-temporal features. Second, KNN and random forest are compared under the same time-ordered train-test split. Third, the numerical results are interpreted through weekday-weekend hourly curves, a weekday-hour heatmap, weather-condition analysis, and permutation importance. The main contribution of this study is to distinguish the predictive roles of temporal and environmental feature groups under a controlled time-ordered experimental design, thereby providing interpretable evidence for feature selection and short-term rebalancing decisions in bike-sharing operations.

2. Methods

2.1. Data source

This study uses the London Bike Sharing Dataset published on Kaggle [10]. The dataset contains 17,414 hourly observations from 2015 to 2017, covering bike-sharing demand, meteorological conditions, and calendar information. The target variable is the hourly number of bike rentals. The average demand in the full dataset is approximately 1,143.10 rentals per hour, which provides a reference for interpreting MAE and RMSE values. The time span includes different seasons, workdays, weekends, holidays, and weather conditions, making it suitable for evaluating the joint contribution of temporal regularity and environmental disturbance.

To avoid programming-style field names in the manuscript, formal names and concise abbreviations are used. Environmental inputs include actual temperature (T1), feels-like temperature (T2), humidity (HUM), wind speed (WS), and weather condition (WC). Temporal inputs include season (Season), holiday indicator (IH), weekend indicator (IW), hour of day (Hour), day of week

DOI: 10.54254/2755-2721/2026.BA36242

(Weekday), and month of year (Month). IH, IW, WS, and WC are used consistently after their first definition.

2.2. Data preprocessing and experimental design

The timestamp was converted to a datetime object, from which Hour, Weekday, and Month were extracted. To mimic the real-world forecasting scenario where only historical observations are available to predict future demand, the dataset was divided chronologically: the first 80% of records were used for training and the last 20% for testing. This produced 13,931 training samples from 2015-01-04 00:00 to 2016-08-10 02:00 and 3,483 test samples from 2016-08-10 03:00 to 2017-01-03 23:00. The same time-ordered split was used for all model and feature-combination experiments to avoid future observations entering the training process.

No missing values were found in the dataset after timestamp parsing and feature extraction; therefore, no imputation or sample deletion was required. Weather condition (WC), Season, holiday indicator (IH), weekend indicator (IW), Weekday, and Month were retained as coded numerical variables. Because KNN is distance-based and sensitive to feature scale, its inputs were standardized using parameters estimated only from the training set, and the fitted transformation was then applied to the test set to prevent data leakage. Random forest is less sensitive to monotonic scale differences and was trained using the original coded values.

Model hyperparameters were tuned exclusively on the training partition using a chronological validation split, with no access to the test set. The training set was further divided chronologically into a training subset and a validation subset: the first 80% of the training set contained 11,144 samples from 2015-01-04 00:00 to 2016-04-15 04:00, and the last 20% contained 2,787 samples from 2016-04-15 05:00 to 2016-08-10 02:00. The validation RMSE was used as the selection criterion. For KNN, the number of neighbors was searched over 3, 5, 7, 10, 15, and 20, with both uniform and distance weighting. The final KNN model used 15 neighbors and distance weighting. For random forest, the number of trees was searched over 50 and 100, the maximum depth over 15, 25, and unlimited, and the minimum number of samples per leaf over 1, 2, and 4. The final random forest used 100 trees, a maximum depth of 25, and a minimum of 2 samples per leaf.

KNN regression was selected as the baseline because of its simplicity, interpretability, and frequent use as a local-distance benchmark, while random forest regression was used as the main model. The random seed of the random forest was set to 42 for reproducibility. Model performance was evaluated using mean absolute error (MAE), root mean square error (RMSE), and the coefficient of determination (R^2). MAE and RMSE are reported in rentals per hour; lower values indicate smaller prediction errors, while an R^2 value closer to 1 indicates stronger explanatory power.

To evaluate the role of feature fusion, three random-forest experiments were conducted with the same tuned parameters: environmental features only, temporal features only, and fused environmental-temporal features. In real operations, future environmental variables would be supplied by weather forecasts. This study uses observed environmental values at the corresponding hours to estimate the incremental predictive value of environmental information under known conditions; therefore, the reported results do not include weather-forecast error.

DOI: 10.54254/2755-2721/2026.BA36242

3. Results and discussion

3.1. Model performance comparison

Table 1. Prediction performance of KNN and random forest

Model	MAE (rentals/h)	RMSE (rentals/h)	R^2
KNN	504.07	765.59	0.540
Random forest	173.08	304.76	0.927

Table 1 shows that random forest clearly outperforms the KNN baseline under the strict time-ordered test setting. The random forest reduces MAE from 504.07 to 173.08 and RMSE from 765.59 to 304.76, corresponding to reductions of 65.7% and 60.2%, respectively. R^2 increases from 0.540 to 0.927, indicating that random forest explains a larger proportion of future demand variation. Compared with a randomly shuffled split, the chronological split avoids leakage from future observations and therefore better reflects practical forecasting difficulty.

This result is consistent with the nonlinear nature of bike-sharing demand. Hourly demand is jointly affected by time, weather, and calendar variables and may follow different rules under different travel scenarios. KNN relies on local distance and can be affected by high-dimensional neighborhood structure, whereas random forest partitions the feature space into multiple demand states and aggregates many decision trees to produce more stable predictions.

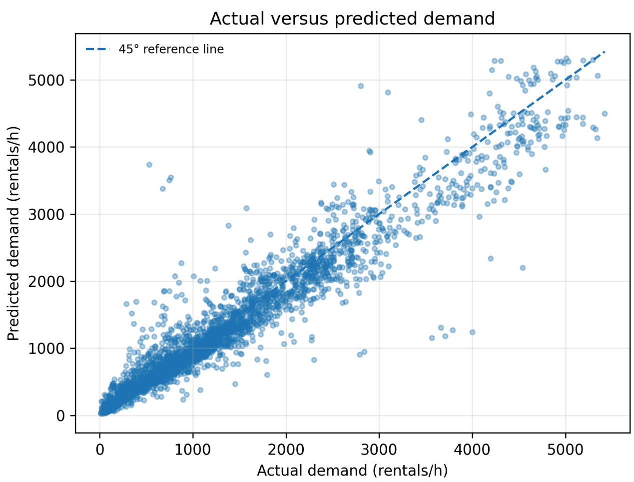

Figure 1. Scatter plot of actual versus predicted hourly demand from the random forest model on the time-ordered test set (photo/picture credit: original)

Most points in Figure 1 are distributed near the 45-degree reference line, which supports the error metrics in Table 1 and indicates that the model captures the overall demand pattern. However, prediction error scales with actual demand magnitude, and high-demand observations tend to show larger dispersion and some underestimation. By demand quartile, the MAE of the highest-demand quartile is approximately 334.03 rentals/h, much higher than the 36.65 rentals/h of the lowest-demand quartile. This indicates that commuting peaks and abnormal demand are still major sources of residual error, consistent with studies showing that extreme weather or public-transport disturbances may move demand away from its regular baseline [7].

3.2. Effect of environmental-temporal feature fusion

Table 2. Random forest performance under different feature combinations

Feature scheme	MAE (rentals/h)	RMSE (rentals/h)	R^2
Environmental	739.32	1051.43	0.132
Temporal	243.63	409.43	0.868
Fused	173.08	304.76	0.927

Note: All feature-combination experiments use the same time-ordered train-test split and tuned random forest hyperparameters.

Table 2 reveals a clear hierarchy among the three feature schemes. Environmental features alone obtain an R^2 of 0.132. This low explanatory power stems from the fact that temperature, humidity, wind speed, and weather condition mainly describe riding comfort and short-term external disturbances but cannot identify stable cycles such as morning and evening peaks, workday commuting, and weekend leisure travel. Temporal features alone achieve an R^2 of 0.868, showing that Hour, Weekday, Month, Season, IH, and IW encode repeated daily and weekly travel rhythms.

The fused feature scheme achieves the best performance, with MAE=173.08, RMSE=304.76, and R^2=0.927. Compared with the temporal-only scheme, feature fusion reduces MAE and RMSE by 29.0% and 25.6%, respectively. This indicates that environmental variables are not strong independent predictors, but they provide meaningful complementary corrections after the temporal demand baseline has been established. For example, demand occurring at the same hour and weekday position can still vary under different temperature, humidity, wind, or weather conditions.

The low R^2 of the environmental-only model should therefore not be interpreted as evidence that weather is unimportant. Instead, it indicates that weather mainly adjusts demand around a time-driven baseline. This conclusion is consistent with Rochas et al., who reported that adding weather information reduced prediction error by more than 20% under degraded weather conditions, although the magnitude of improvement differs because of different model structures, datasets, and weather distributions [8].

3.3. Temporal demand pattern analysis

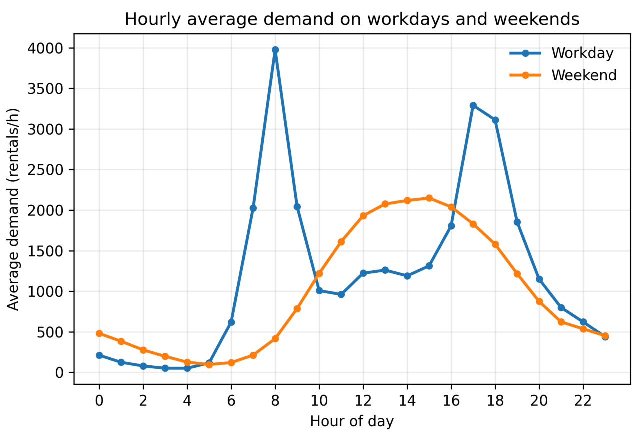

Figure 2. Average hourly demand profiles for workdays (non-holiday weekdays) and weekends (photo/picture credit: original)

DOI: 10.54254/2755-2721/2026.BA36242

Figure 2 explains why temporal features have strong predictive power in Table 2. To avoid treating statutory holidays as ordinary workdays, this study defines workdays as IW=0 and IH=0, while weekends are defined as IW=1. The 384 hourly holiday records falling on non-weekend dates are excluded from this curve comparison. Workday demand reaches its highest point at 08:00, with an average of approximately 3,977 rentals, and forms a second high-demand interval at 17:00 and 18:00, with average values of approximately 3,290 and 3,112 rentals, respectively. In contrast, the weekend curve is smoother, reaches its peak later, and records its maximum average demand of approximately 2,148 rentals at 15:00.

Therefore, IW and IH help distinguish workdays, weekends, and holidays and reflect different intraday demand-curve shapes. For practical operations, routine rebalancing schedules can be primarily designed according to temporal periodic patterns, while weather forecasts can be used to fine-tune dispatch volumes during peak hours.

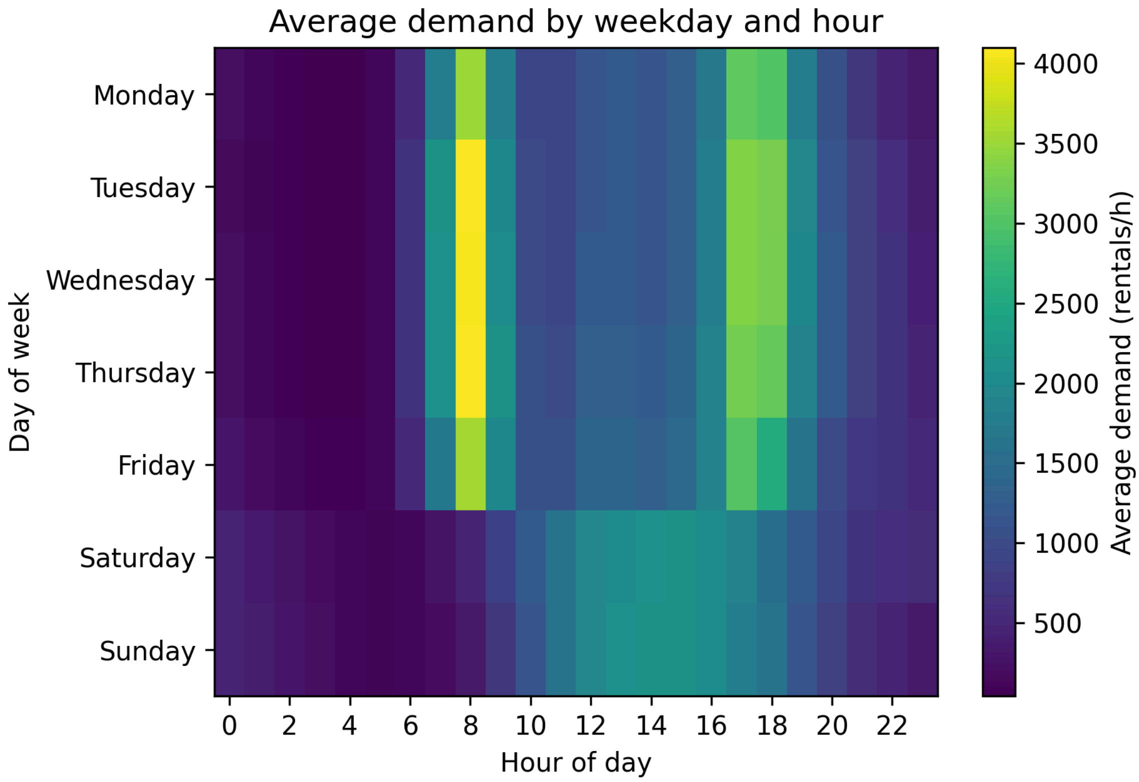

Figure 3. Heatmap of average hourly demand by weekday and hour (photo/picture credit: original)

Figure 3 further visualizes demand patterns along the two temporal dimensions of Weekday and Hour. The heatmap shows repeated high-demand bands in the morning and evening from Monday to Friday, while Saturday and Sunday display a more dispersed distribution without the same sharp commuting peaks. This pattern demonstrates that hourly bike-sharing demand is not random but follows a stable daily and weekly structure.

The visual evidence in Figures 2 and 3 is consistent with the temporal-only R^2 of 0.868 in Table 2. Temporal features are effective because they encode recurring activity schedules such as work trips, school trips, and non-commuting travel. However, Table 2 also shows that relying only on temporal features is not optimal: adding environmental variables reduces MAE and RMSE by 29.0% and 25.6%, respectively.

DOI: 10.54254/2755-2721/2026.BA36242

3.4. Environmental factors and feature importance analysis

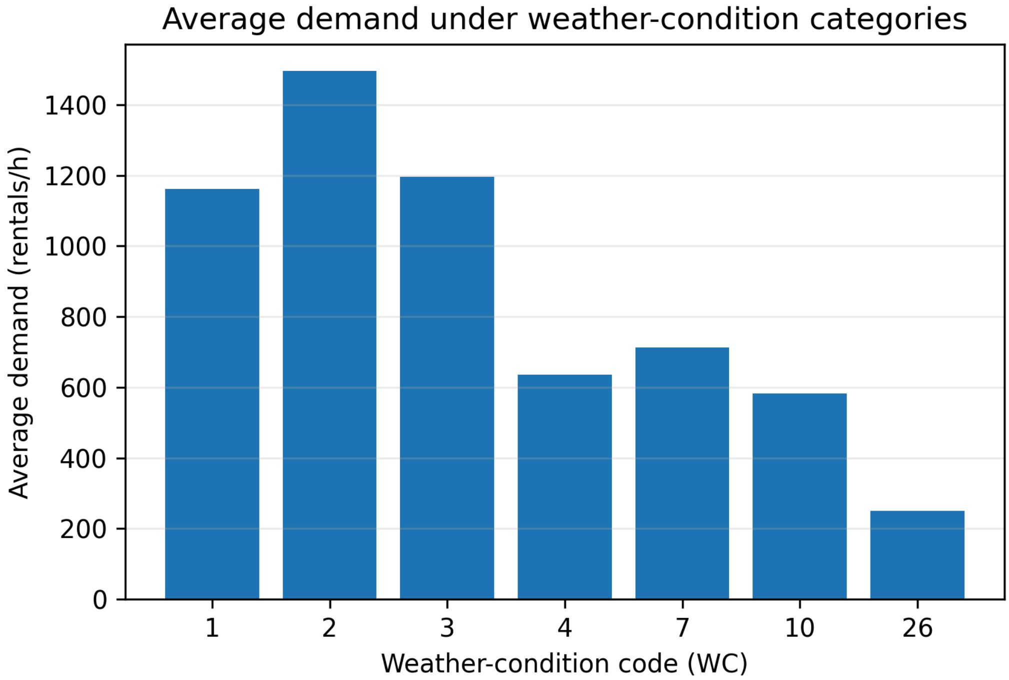

Figure 4. Average hourly demand under different weather-condition categories (photo/picture credit: original)

Figure 4 provides direct descriptive evidence for the supplementary role of environmental variables. According to the dataset coding, WC=1 represents clear weather, WC=2 scattered clouds, WC=3 broken clouds, WC=4 cloudy conditions, WC=7 light rain, WC=10 heavy rain, and WC=26 snowfall. Average demand differs considerably across categories. WC=2 has the highest average demand at 1496.18 rentals, whereas WC=26 has the lowest average demand at 250.85 rentals. For descriptive comparison, WC=1-3 are grouped as relatively mild weather and WC=4, 7, 10, and 26 as relatively adverse weather. The two groups have average demands of 1268.75 and 674.00 rentals, respectively, representing a relative decrease of 46.9%. This grouping is used only for descriptive comparison and is not an official dataset classification.

This result should be interpreted as an association rather than a causal effect. Some rare weather categories have limited observations; for example, WC=10 and WC=26 contain only 14 and 60 records, respectively. Despite this limitation, the overall pattern supports the feature-fusion result in Table 2: environmental conditions do not independently determine hourly demand, but they help correct the temporal demand baseline.

Figure 5 uses 20 repeated permutations to measure each variable's contribution to prediction performance on the time-ordered test set. Error bars represent the standard deviation of the 20 permutation results. The metric is the average decrease in test-set R^2 after one feature is shuffled, rather than a normalized percentage. When shuffling an essential feature reduces model R^2 below zero, its importance value can exceed 1. The results show that Hour has a much larger importance than all other variables. Weekday and IW rank next, indicating that weekday position and workday-weekend differences jointly shape the intraday demand curve. IH also contributes to prediction, while Season and Month have relatively low single-variable contributions.

DOI: 10.54254/2755-2721/2026.BA36242

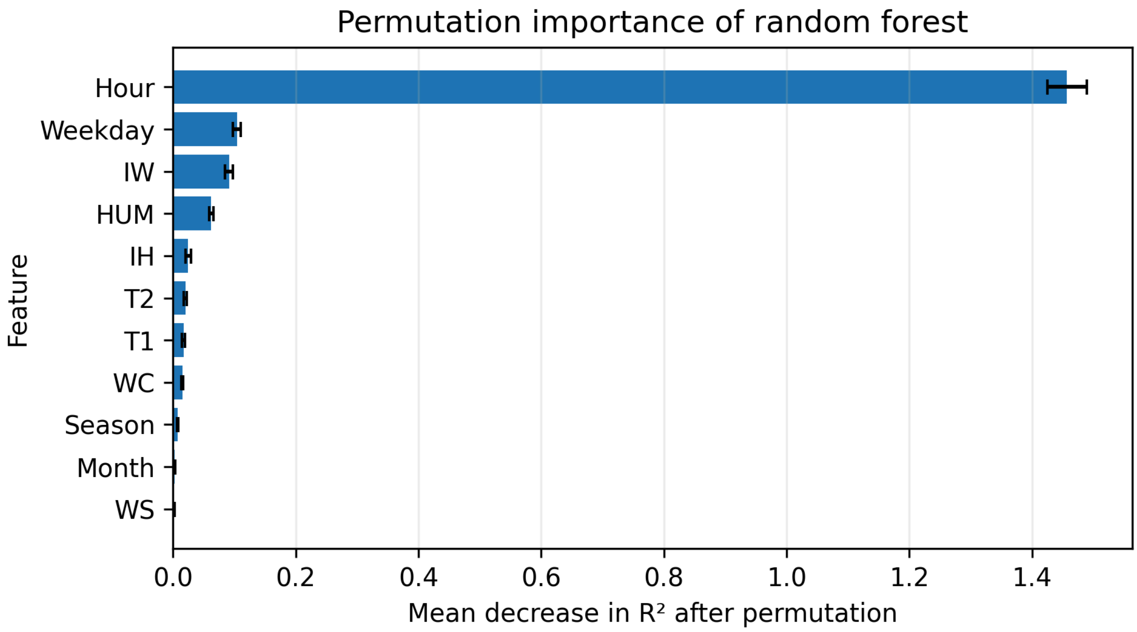

Figure 5. Permutation-importance ranking of the random forest model with standard-deviation error bars (photo/picture credit: original)

Among environmental variables, HUM has the highest permutation importance, followed by T2, T1, WC, and WS, suggesting that humidity and temperature-related conditions provide corrective information for the time-driven baseline. Because T1 and T2 both describe thermal conditions, their high correlation means that shuffling one variable may leave part of the temperature information in the other variable, so single-variable importance can underestimate their joint information value. Therefore, feature-importance rankings should be interpreted as model-specific evidence rather than causal strength, and they should be considered together with the feature-combination experiment in Table 2.

It should also be noted that feature importance is an explanatory diagnostic rather than causal proof. Correlated predictors may substitute for each other, and the demand structure may vary across time periods. Future studies can use time-segmented permutation tests, SHAP values, or performance changes after removing correlated variables to examine the stability of these conclusions. Nevertheless, the evidence from Table 2 and Figures 2-5 is internally consistent: temporal regularity explains most demand variation, while environmental information further improves future-period prediction by correcting the temporal baseline.

4. Conclusion

This study investigates hourly bike-sharing demand prediction using 17,414 records from the London Bike Sharing Dataset. Under a chronological 80:20 train-test split, the tuned random forest achieves MAE=173.08, RMSE=304.76, and R^2=0.927, reducing MAE and RMSE by 65.7% and 60.2% compared with KNN. Feature-combination experiments further show that environmental features alone have limited explanatory power (R^2=0.132), temporal features capture most of the demand structure (R^2=0.868), and fused environmental-temporal features provide the best performance. Relative to the temporal-only setting, the fused scheme reduces MAE and RMSE by 29.0% and 25.6%, respectively.

DOI: 10.54254/2755-2721/2026.BA36242

The visual analyses and permutation-importance results support the same conclusion. Workday-weekend hourly curves and the weekday-hour heatmap reveal stable commuting-related patterns, while weather-condition analysis shows that adverse weather categories are associated with lower average demand. Permutation importance confirms that Hour is the dominant predictor, with environmental variables acting as supplementary corrections. Overall, temporal features establish the main demand baseline, and environmental features refine that baseline under changing riding conditions.

This study has several limitations. First, it uses an aggregated single-city dataset and does not include station-level spatial heterogeneity. Second, environmental inputs are based on observed values, so weather-forecast errors are not included in end-to-end prediction. Third, the analysis focuses on one-step hourly prediction and does not evaluate multi-step horizons such as 6-hour or 24-hour-ahead forecasting. Finally, the dataset covers 2015-2017, and long-term changes in system scale, user behavior, and travel habits may introduce trend drift that is not explicitly modeled. Future work can incorporate station-level spatial variables, event and public-transport-disruption data, real weather forecasts, rolling prediction updates, spatiotemporal models, SHAP values, and time-segmented permutation tests to further evaluate generalizability and interpretability.

References

[1] Zhou, X. (2015). Understanding spatiotemporal patterns of biking behavior by analyzing massive bike sharing data in Chicago. *PLOS ONE, 10*(10), Article e0137922. https://doi.org/10.1371/journal.pone.0137922

[2] Jiao, Z. L., Jin, H., Liu, B. L., et al. (2018). Short-term demand prediction of shared bicycles driven by big data: A comparative analysis based on machine-learning models. *Journal of Business Economics*, (8), 16–25, 35. https://doi.org/10.14134/j.cnki.cn33-1336/f.2018.08.002

[3] El-Assi, W., Mahmoud, M. S., & Habib, K. N. (2017). Effects of built environment and weather on bike sharing demand: A station level analysis of commercial bike sharing in Toronto. *Transportation, 44*(3), 589–613. https://doi.org/10.1007/s11116-015-9669-z

[4] Xie, G. W., Qian, L. B., & Tang, W. Y. (2022). Relationship between weather and urban bike-sharing travel demand. *Logistics Sci-Tech*, (4), 84–90. https://doi.org/10.13714/j.cnki.1002-3100.2022.04.015

[5] Hu, Y. Q., Hamiti, & Xu, Z. K. (2023). Demand prediction for shared bicycles around urban rail stations based on K-means clustering. *Journal of Transport Engineering, 23*(5), 82–90. https://doi.org/10.13986/j.cnki.jote.2023.05.014

[6] Li, X., Xu, Y., Zhang, X., Shi, W., Yue, Y., & Li, Q. (2023). Improving short-term bike sharing demand forecast through an irregular convolutional neural network. *Transportation Research Part C: Emerging Technologies, 147*, Article 103984. https://doi.org/10.1016/j.trc.2022.103984

[7] Rennie, N., Cleophas, C., Sykulski, A. M., & Dost, F. (2023). Analysing and visualising bike-sharing demand with outliers. *Discover Data, 1*, Article 1. https://doi.org/10.1007/s44248-023-00001-z

[8] Rochas, R., Furno, A., & El Faouzi, N.-E. (2023). Contextual data integration for bike-sharing demand prediction with graph neural networks in degraded weather conditions. In *2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC)* (pp. 5436–5441). IEEE. https://doi.org/10.1109/ITSC57777.2023.10422448

[9] Behroozi, A., & Edrisi, A. (2025). Predicting travel demand of a bike sharing system using graph convolutional neural networks. *Public Transport, 17*, 281–317. https://doi.org/10.1007/s12469-024-00371-w

[10] Mavrodiev, H. (2019). London bike sharing dataset [Data set]. *Kaggle*. https://www.kaggle.com/datasets/hmavrodiev/london-bike-sharing-dataset