

Predicting bicycle demand at the station level by incorporating station-specific spatial features in graph convolutional network model

Jonghyun Choi^a, Kyungtaek Oh^a, Seungmin Jang^a, and Keungoui Kim^{*a}

^aSchool of Applied Artificial Intelligence, Handong Global University, Pohang, South Korea

ARTICLE HISTORY

Compiled March 15, 2024

ABSTRACT

Shared bicycle rental services have achieved widespread success, particularly in areas with heavy traffic. However, imbalanced and irregular usage patterns have emerged as significant issues. In the context of shared bicycle rental services, stations for renting and returning bicycles differ, creating a network that can be modeled as a graph with stations as nodes and user interactions as edges. In this regard, the Graph Convolutional Network (GCN), a deep learning model adept at capturing the relational structure of networks, has been introduced as a method for predicting bicycle usage and weather data. This research proposes an advanced GCN model that incorporates spatial features and optimized hyperparameters for predicting daily bicycle rental volumes at the station level. Our model surpasses traditional forecasting models and other GCN models that do not explicitly account for station-level spatial characteristics. Accurately predicting bicycle demand at the station level has been challenging due to the intricate interplay of relationships and spatio-temporal dynamics. By including precise spatial features specific to each station, our model offers a more detailed representation of the complex network essential for demand prediction, resulting in enhanced prediction accuracy.

KEYWORDS

Bicycle Demand Prediction, Bicycle Rental Network, Graph Convolution Network, Spatial Feature Matrix, Shared-bicycle Rental Service

1. Introduction

The growing preference for environmentally friendly and healthy lifestyles has led to a continuous increase in bicycle demand (Agarwal, Ziemke, & Nagel, 2020). Alongside the explosive growth of shared mobility services, bicycle sharing service has emerged in various cities as an efficient and sustainable mode of transportation (Li, Fuerst, & Luca, 2023). As the public becomes more aware of the numerous benefits of using bicycles—such as time savings, cost efficiency, and health benefits, bicycle-sharing services have started to be recognized as a viable alternative mode of transport (Cui, 2018). It is particularly true in densely populated and traffic-congested cities like Seoul, where bicycle-sharing services offer a sustainable and green transportation solution (Park, Honda, Fujii, & Kim, 2022).

Regardless of its effectiveness, bicycle-sharing services face a significant challenge in managing supply and demand due to varying utilization rates across different stations (Singhvi et al., 2015). Stations with high usage often experience a shortage of available bicycles, whereas stations with lower usage grapple with an excess of parked bicycles (Kloimüllner, Papazek, Hu, & Raidl, 2014). In this sense, it is crucial to forecast bicycle demand at docking stations by analyzing usage patterns (P. Chen et al., 2020; Xia, Jiang, & Wang, 2022). With the help of accurate prediction, the number of bicycles per station can be efficiently managed, and flexible

bicycle distribution can be implemented.

The prediction of bicycle demand in bicycle-sharing services is, however, a highly complex problem to solve (Y. Zhou, Wang, Zhong, & Tan, 2018). For instance, the demand for shared bicycles can be influenced by various factors such as geographical location, time, weather conditions, mobility conditions, movement patterns, socio-economic features, etc (H. Xu, Ying, Wu, & Lin, 2013). This results in the unpredictable nature of user preferences, making it much more challenging to identify usage patterns among heterogeneous stations. Consequently, the field of bicycle-sharing system demand prediction has been diverse research efforts focused on understanding and forecasting user patterns (Boufidis, Nikiforidis, Chrysostomou, & Aifadopoulou, 2020; Fanaee-T & Gama, 2014; Gebhart & Noland, 2014; Giot & Cherrier, 2015; Salaken, Hosen, Khosravi, & Nahavandi, 2015; V E & Cho, 2020). From this perspective, a deep-learning-based approach can offer a novel solution to handle this complex problem (Cagliero, Cerquitelli, Chiusano, Garza, & Xiao, 2017; X. Yang & He, 2020). When predicting with a deep-learning model, one important aspect to consider carefully is the choice of the deep-learning model, ensuring that its advantages and unique features align well with the problem at hand. The demand and supply of bicycle-sharing services are like a problem of a network, and this implies that a Graph Convolutional Network (GCN) can be implemented.

In previous studies, the GCN model, a deep learning model adept at capturing the relational structure of networks, has been introduced as a prediction method (Chai, Wang, & Yang, 2018; R. Guo et al., 2019; S. Guo, Lin, Feng, Song, & Wan, 2019; Kim, Lee, & Sohn, 2022; L. Lin, He, & Peeta, 2018; Zi, Xiong, Chen, & Chen, 2021). Formerly introduced GCN models were primarily focused on incorporating the relationships between bike rental stations into the graph. These models have demonstrated the effectiveness of reflecting the rental history, distance, and correlations between bike stations as spatial characteristics in the adjacency matrix. However, since individual predictions for each bike rental station are crucial for bike redistribution, there is a growing emphasis on capturing the dynamic spatial characteristics of individual bike stations. Nevertheless, the previously reported GCN models have limitations, as they only reflect spatial features in the edges, which represent node relationships in the graph data. Consequently, they fall short of adequately explaining the specific spatial characteristics of individual rental stations.

In this paper, we developed a Graph Convolutional Network pro (GCN-pro) model to predict demand at each rental station. To provide a more detailed description of the intricate patterns at individual bike stations, we defined the spatial characteristics of each station based on its proximity to nearby transportation facilities, residential areas, cultural establishments, educational facilities, and other relevant features. This approach was inspired by numerous prior studies reporting a strong correlation between the willingness to use the sharing-bike system and the spatial accessibility factors (Du, Deng, & Liao, 2019; P. Lin et al., 2020; L. Yang et al., 2023; H. Zhang, Zhuge, Jia, Shi, & Wang, 2021). Additionally, we defined the relationships between bike stations using historical rental and return records, considering stations with more extensive rental and return histories as having higher relational significance. Furthermore, by processing daily and weekly data separately and performing operations on adjacency matrices, we identified temporal patterns on a daily and weekly basis. Finally, we incorporated weather variables into the analysis.

2. Literature Review

Predicting bicycle demand requires a thorough understanding of complex spatiotemporal factors, which was a challenging task before the advancement of deep learning. This complexity has prompted researchers to explore cluster-based approaches, wherein stations with similar

characteristics are grouped, simplifying predictions by forecasting demand at the group level (L. Chen et al., 2016; Feng, Chen, Du, Li, & Jing, 2018; Liu et al., 2017). Clustering operates based on usage patterns (L. Chen et al., 2016) and geographical characteristics (Feng et al., 2018; Liu et al., 2017). However, with the exacerbation of the bicycle redistribution problem, the importance of accurate station-level demand prediction has been underscored, necessitating a more detailed analysis of both spatial and temporal factors. Therefore, several attempts have been made, as evidenced by the works of Hulot, Aloise, and Jena (2018); Ramesh, Nagisetti, Sridhar, Avery, and Bein (2021); C. Xu, Ji, and Liu (2018); H. Xu, Duan, and Pu (2019); T. Xu et al. (2020); Zhao, Zhao, Xia, and Jia (2022). They initially approached the problem by examining major traffic behaviors and employing simple models such as XGBoost and linear regression (Hulot et al., 2018; Ramesh et al., 2021). Subsequently, they played a crucial role in analyzing travel patterns and predicting demand more accurately by utilizing techniques that consider spatial relationships between stations, such as LSTM neural networks, K-means clustering, self-organizing maps, and hyper-clustering (C. Xu et al., 2018; H. Xu et al., 2019; T. Xu et al., 2020; Zhao et al., 2022).

Incorporating the pioneering work of Kipf and Welling (2017) on Graph Convolutional Networks (GCN) for graph data processing, subsequent research has sought to develop bicycle-sharing demand prediction by melding spatial with temporal insights. By encoding spatial characteristics between bike stations in the adjacency matrix and learning past bike rental patterns through LSTM, spatial-temporal factors were strengthened (Chai et al., 2018; L. Lin et al., 2018). Furthermore, attention mechanisms were introduced to capture real-time correlations, and fully connected networks were utilized to predict demand (S. Guo et al., 2019; Kim et al., 2022; Zi et al., 2021). However, the nuanced spatial characteristics of individual stations remain largely unexplored, with a predominant focus on temporal data. This ongoing journey underscores significant advancements for improvement in spatial-temporal predictive modeling within the bicycle-sharing domain, emphasizing the need for balanced consideration of spatial features and temporal dynamics.

3. Materials and Methods

3.1. Framework: GCN-pro

The general framework of the GCN model typically comprises nodes and edges, representing relationships between nodes (S. Zhang, Tong, Xu, & Maciejewski, 2019). In this study, bicycle stations are designated as nodes, and the relationships between stations are defined as edges (L. Lin et al., 2018). Specifically, the relationships between stations are characterized as the connectivity between bicycle rental and return stations, utilizing historical data on rentals and returns. The overall structure of our proposed model is derived from the GCN model proposed by Kim et al. (2022), with certain modifications. Here, authors separately used historical rental patterns on an hourly, daily, and weekly basis, and then trained a fully connected network along with weather variables. The presented model, however, fails to account for the complex spatial attributes unique to each station, resulting in inadequate attention to the spatial dependencies that vary from one station to another.

In contrast, our proposed model integrates spatial characteristics specifically tailored to individual rental stations. To the best of our knowledge, such an approach has not been addressed in prior literature. A graph is designed as $G = (N, E, s_i, r_i)$, where N is the set of stations, E is the set of edges, s_i represents the spatial features of every station i , and r_i denotes the actual rental records of every station i . Our GCN model uniquely captures the spatial attributes of each station alongside the daily and weekly rental patterns, aiming for accurate predictions for upcoming dates (as depicted in Figure 1). In this model, the nodes symbolize bicycle rental stations, while the edges are characterized by historical bicycle rental and return records.

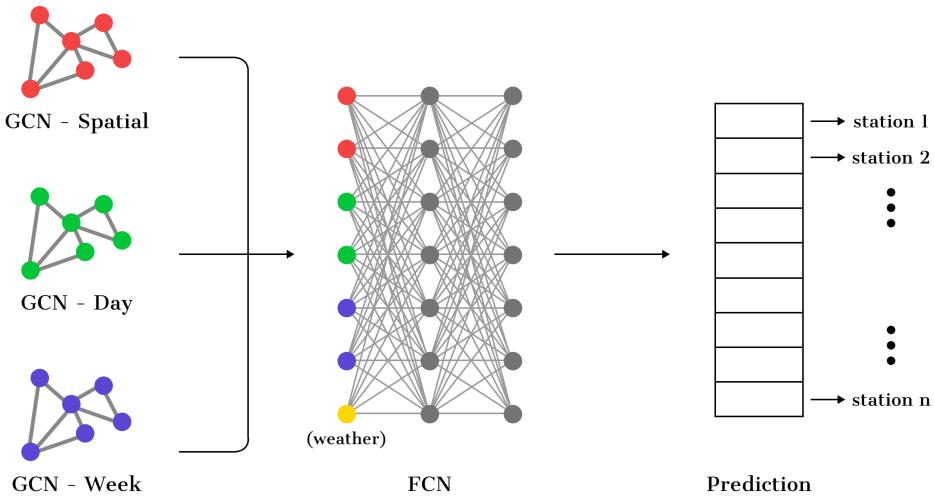


Figure 1. Schema of the GCN-pro

To construct a comprehensive model that assimilates all these features, we integrate the outputs of the GCN models into a Fully Connected Network (FCN). This is achieved by concatenating the outputs from each GCN model into a single-dimensional vector, which is then fed into the first layer of the FCN. Moreover, weather variables, illustrated as yellow dots in Figure 1, are incorporated to predict daily rental demand for each station. This model is adept at learning from the spatial characteristics of stations, historical bicycle usage data, and weather conditions, such as precipitation and temperature, on the prediction date. The spatial features specific to each station reflect their proximity to key amenities like subway stations, bus stops, residential areas, and bicycle paths. We define proximity as the number of relevant spatial features within a 500-meter radius of the rental stations. Through the Kruskal-Wallis test, we identify variables within these spatial features that significantly influence rental demand. Only these significant features are integrated into our model, thereby enhancing its predictive accuracy. We term this model ‘GCN-pro’.

3.2. Data Preparation

For the empirical analysis, data on Seoul’s daily bicycle usage, spatial information, and weather conditions were utilized. The city has seen the successful implementation and widespread adoption of a public bicycle-sharing service known as ‘Ddareungi’¹. Our first dataset, the daily bicycle usage dataset, spanning from March to September 2022 for Seoul’s public bike rental stations, was sourced from ‘Seoul Open Data Plaza’. It includes ride records that provide insights into rental and return stations, and usage dates, amassing a total of 868,451 rental records across 2,763 bike stations.

Significant pre-processing was required to prepare the dataset for analysis. This process involved removing records with zero travel distance, which were considered missing values based on recommendations from the Seoul Metropolitan Government’s Department of Transportation. Further filtering was applied to ensure data accuracy and relevance, leading to the exclusion

¹<https://www.bikeseoul.com>

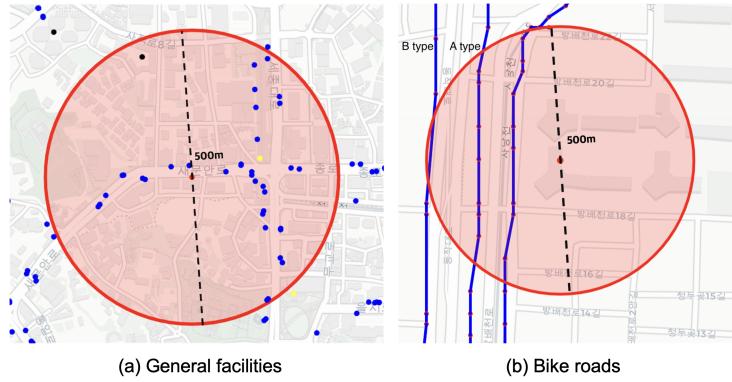


Figure 2. Counting the number of general facilities and bike roads near bicycle stations

(a) The provided illustration depicts the process of quantifying the number of general facilities in the vicinity of a single bicycle station. The red dot represent bicycle stations, the blue dots denote bus stops, the yellow dots signify subway stations, and the black dots represent residential facilities. The counting methodology involves establishing a 500m radius around each station, within which the number of facilities is tallied. Consequently, for the highlighted station, the count reveals 35 bus stops, 1 subway station, and a number of households in the residential facility. (b) Two scenarios were considered: in the A type, roads were counted if at least one bending point was within a 500-meter radius to prevent duplicates. In the B type, when all bending points were outside the 500-meter radius, linear equations and the distance from the bicycle station to the intersection point were used to count roads that were less than 500 meters. Through this preprocessing, spatial data consisting of 15 variables within a 500-meter radius for each bicycle rental station was generated.

of 54 stations due to inaccurate location data, 66 stations that were not used in 2022, and 88 stations impacted by mid-year closures or openings. The final summary of the bicycle docking stations and rental records used in the study is presented in Table 1. The final dataset was divided into a training set (March to August) and a test set (September), ready for further examination.

Period	Number of stations	Rental records
March, 2022 - September, 2022	2,555	864,980

Table 1. Summary of data sample

Secondly, spatial data was obtained from the ‘Seoul Open Data Plaza’ and ‘Smart Seoul Map’. Spatial data was employed to extract information about facilities within a 500-meter radius of Seoul bicycle stations (Q. Zhou, Wong, & Su, 2020). Data downloaded from ‘Seoul Open Data Plaza’ and ‘Smart Seoul Map’ included information about bus stops, subway stations, schools, universities, apartments, cultural facilities, and large academies. Spatial data pre-processing was performed to determine the count of these facilities around each bicycle station (Figure 2(a)). Buildings that closed or were demolished before 2022 and those that were constructed or opened after 2022 were excluded.

Bike road variables were divided into (A) Mixed-Use Road (Non-Separated), (A) Exclusive Bicycle Lane, (A) Separated Mixed-Use Road, (A) Exclusive Bicycle Lane for Vehicles, (B) Exclusive Bicycle Lane, (B) Separated Mixed-Use Road, (B) Mixed-Use Road (Non-Separated), and (A) Bicycle Priority Road. Regarding the processing of bicycle roads in the spatial data, roads with a radius of 500 meters centered around bicycle stations were formed by connecting bending points (Figure 2(b)).

Lastly, weather data, crucial for demand prediction, was acquired from the ‘Korea Meteorological Administration’, encompassing nine variables such as precipitation (mm), maximum temperature (°C), average temperature (°C), minimum temperature (°C), diurnal temperature

range, average humidity, sunshine duration, average wind speed (m/s), and maximum wind speed (m/s). Although the precipitation variable had numerous missing values, the missing values were handled by considering them as the absence of rain and replacing NA values with 0. This pre-processing resulted in a time-series format containing nine weather variables spanning from March 1, 2022, to September 33, 2022.

3.3. Spectral graph convolutions

The utilization of the GCN model serves as a fundamental tool in addressing complex traffic issues (Yu, Yin, & Zhu, 2018). The decision to adopt GCN is rooted in its distinct capability to capture and analyze the intricate relationships between stations present in bicycle rental systems. To effectively predict bicycle demand, the GCN model proposed by Kipf and Welling (2017) is employed. The model is grounded in Spectral Graph Convolution. Through the eigenvalue decomposition of the graph Laplacian matrix L , the graph is transformed into the domain of bicycle demand prediction. The matrix U , comprised of eigenvectors, enables an effective representation of graph characteristics.

$$g_{\theta} \cdot x = U g_{\theta} U^{\top} x \quad (1)$$

In Eq. (1), x is a scalar associated with each node, and the function g_{θ} represents a filter applied to the graph signal x or a function that emphasizes or transforms features of the graph signal. θ denotes the parameters of the filter, and U represents the matrix of eigenvectors derived from the normalized graph Laplacian $L = I_N - D^{-\frac{1}{2}} A D^{-\frac{1}{2}} = U \Lambda U^{\top}$, where Λ is the diagonal matrix of eigenvalues, and $U^{\top} x$ denotes the graph transform of x .

Due to the computational cost of matrix multiplication with U , however, an approximation using Chebyshev polynomials was introduced following the proposal by Hammond, Vandergheynst, and Gribonval (2011). This approach approximates the graph Laplacian with a K -th degree polynomial, providing K -localized features that depend only on nodes within K steps. Here, restricting K to 1 allows for the stacking of multiple graph convolution layers to form a deep model, while linearly constraining each layer concerning the Laplacian enables the utilization of diverse filter function classes (Kipf & Welling, 2017). Consequently, the expression representing the result of graph convolution operations is ultimately summarized in Eq. (2).

$$Z = \tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}} X \Theta \quad (2)$$

where $\tilde{A} = A + I_N$. Here, \tilde{A} is a matrix obtained by adding the identity matrix to the adjacency matrix, incorporating self-loops for each node. Furthermore, $\tilde{D}_{ii} = \sum_j \tilde{A}_{ij}$. Additionally, $\Theta \in \mathbb{R}^{C \times F}$ is a matrix of filter parameters, and $Z \in \mathbb{R}^{N \times F}$ is the convolved signal matrix, where C represents the dimensionality of node features, F denotes the number of feature maps, and N signifies the number of nodes.

3.4. Adjacency Matrix: Interaction among stations

The adjacency matrix demonstrates the relationships between bicycle stations, using total historical data for bike rentals and returns. This matrix is arranged such that both rows and columns correspond to station numbers, ensuring they are identical. Rows indicate the numbers of the stations where bikes are rented, and columns show the numbers of the stations where bikes are returned.

$$A = \begin{bmatrix} R_0R_0 & R_0R_1 & R_0R_2 & \dots & \dots & \dots & R_0R_N \\ R_1R_0 & R_1R_1 & R_1R_2 & \dots & \dots & \dots & R_1R_N \\ R_2R_0 & R_2R_1 & R_2R_2 & \dots & \dots & \dots & R_2R_N \\ \dots & \dots & \dots & \dots & R_iR_j & \dots & \dots \\ R_NR_0 & R_NR_1 & R_NR_2 & \dots & \dots & \dots & R_NR_N \end{bmatrix} \quad (3)$$

Here, R_iR_j represents the total count of bicycles moved from station i to station j . Consequently, the adjacency matrix A encapsulates information regarding interactions between different bicycle stations, representing the records of all bicycles in terms of their origin and destination stations. However, since this matrix encompasses the frequency of all bicycle rental histories, its values were scaled using min-max scaling to bring them within the range of 0 to 1. This adjustment facilitates direct operations with the feature matrix.

$$L^{\text{sym}} = \tilde{D}^{(-1/2)} \tilde{A} \tilde{D}^{(-1/2)} \quad (4)$$

In Eq. (4), \tilde{D} is the diagonal matrix derived from the matrix \tilde{A} , and L^{sym} is the normalized symmetric Laplacian Matrix obtained by multiplying both sides of the scaled adjacency matrix with the power of -1/2 of the diagonal matrix. The construction of a symmetric Laplacian Matrix serves the purpose of efficient information management when performing operations with the feature matrix in eigenvalue-eigenvector forms (Kipf & Welling, 2017).

3.5. Feature Matrix 1: spatial feature by station level

As mentioned earlier, this research introduces a novel approach to constructing spatial matrices that have not been explored previously. It includes the spatial characteristics specific to individual bicycle stations, encapsulating parameters such as the number of transportation facilities, residential households, and other related features within a 500-meter radius of each station.

3.5.1. Examining the Impact of Spatial Features on Bicycle Rental Quantities

A thorough investigation was undertaken to determine the influence of 15 spatial features on actual rental quantities. At the study's outset, groups were categorized based on quartiles related to the number of spatial features, creating four distinct groups. The null hypothesis proposed that "there is no significant difference in rental quantities among groups distinguished by varying numbers of spatial features." The non-parametric Kruskal-Wallis test was employed to evaluate this hypothesis, a choice driven by its suitability for scenarios lacking homogeneity of variance among examined groups, and its ability to detect differences in means across three or more groups. Given the violation of the homoscedasticity assumption by the groups in this study, the Kruskal-Wallis test was selected. The significance threshold for the Kruskal-Wallis test was set at 0.05. Additionally, Levene's test was performed to verify the homogeneity of variance assumption, with a significance level of 0.05 applied to assess homoscedasticity. Upon establishing statistical significance, modeling proceeded by including only the spatial features identified as impactful. This approach was informed by the interpretation that these features significantly influenced bicycle rental quantities, encompassing a total of 11 characteristics (Table 2).

3.5.2. Spatial matrix

A spatial matrix is constructed as one of the feature matrices that will be operated on with the adjacency matrix. This spatial matrix is formed based on eleven spatial features that have had a substantial impact on the actual rental quantities. The rows of the matrix correspond to a total of unique bicycle stations, while the columns represent 11 distinct spatial characteristics.

Spatial feature	p-value
Bus station	7.72e-18
Train station	7.74e-15
School	0.01
Apartment household	0.04
Cultural facility	0.30
University	0.83
Academey	0.18
(A) Mixed-Use Road (Non-Separated)	4.66e-08
(A) Exclusive Bicycle Lane	1.86e-24
(A) Separated Mixed-Use Road	0.0003
(A) Exclusive Bicycle Lane for Vehicles	6.15e-05
(B) Exclusive Bicycle Lane	1.55e-05
(B) Separated Mixed-Use Road	0.001
(B) Mixed-Use Road (Non-Separated)	0.46
(A) Bicycle Priority Road	7.29e-10

Table 2. Kruskal-Wallis test result

The presented p-values in the table are derived from Kruskal-Wallis tests conducted on 15 spatial features considered in the paper. The statistical results indicate that 11 specific characteristics, including the 'number of nearby bus stops,' 'number of nearby subway stations,' 'number of nearby schools,' (A) Mixed-Use Road (Non-Separated), (A) Exclusive Bicycle Lane, (A) Separated Mixed-Use Road, (A) Exclusive Bicycle Lane for Vehicles, (B) Exclusive Bicycle Lane, (B) Separated Mixed-Use Road, (B) Mixed-Use Road (Non-Separated), and (A) Bicycle Priority Road, significantly influence the bicycle rental quantities.

$$F_{\text{spatial}} = \begin{pmatrix} S_1 & S_2 & \dots & S_{11} \\ x_{11} & x_{12} & \dots & x_{1,11} \\ x_{21} & x_{22} & \dots & x_{2,11} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{n,11} \end{pmatrix} \quad (5)$$

where S_j denotes spatial features, and $x_{i,j}$ represents the count of spatial feature j near the i th station. The matrix is transformed into values between 0 and 1 using min-max scaling, aiming to prevent the model from biased computations towards specific features.

3.6. Feature Matrix 2: temporal pattern

One of the critical attributes in predicting bicycle demand is understanding the temporal patterns of previous bicycle rental volumes. To achieve this, temporal matrices are constructed to train the model on the historical daily and weekly bicycle rental patterns for each station. These matrices are categorized into daily and weekly temporal matrices, representing the rental history for several days and weeks before the prediction date, respectively.

$$F_{\text{day}} = \begin{pmatrix} t-d & \dots & t-2 & t-1 \\ R_{1,t-d} & \dots & R_{1,t-2} & R_{1,t-1} \\ R_{2,t-d} & \dots & R_{2,t-2} & R_{2,t-1} \\ \vdots & \vdots & \ddots & \vdots \\ R_{n,t-d} & \dots & R_{n,t-2} & R_{n,t-1} \end{pmatrix} \quad (6)$$

$$F_{\text{week}} = \begin{pmatrix} t-w & \dots & t-2 & t-1 \\ R_{1,t-w} & \dots & R_{1,t-2} & R_{1,t-1} \\ R_{2,t-w} & \dots & R_{2,t-2} & R_{2,t-1} \\ \vdots & \vdots & \ddots & \vdots \\ R_{n,t-w} & \dots & R_{n,t-2} & R_{n,t-1} \end{pmatrix} \quad (7)$$

Equation 6 represents the daily historical rental volume, while Equation 7 represents the weekly historical rental volume. Here, ‘t’ denotes the day under prediction, while ‘n’ represents the total number of stations. In Equation 6, $R_{i,t-d}$ indicates the rental volume at the i-th station d days before the prediction date, while in Equation 7, $R_{i,t-w}$ represents the total rental volume of the i-th station w weeks before the prediction date. The values of d and w indicated in the matrices have been set as hyperparameters.

3.7. GCN-pro propagation

In this chapter, the process of propagation of GCN-pro will be discussed, delving into how the matrices created earlier are used for training within the model.

$$\begin{aligned} H_1^{(0)} &= F_{\text{spatial}} \\ H_2^{(0)} &= F_{\text{day}} \\ H_3^{(0)} &= F_{\text{week}} \end{aligned} \quad (8)$$

Here, F_{spatial} is designated as the first layer of matrix H_1 , F_{day} as the first layer of matrix H_2 , and F_{week} as the first layer of matrix H_3 . Each feature matrix serves as an individual input to the GCN model.

$$\begin{aligned} H_1^{(l+1)} &= \sigma(\hat{D}^{(-1/2)} \hat{A} \hat{D}^{(-1/2)} H_1^{(l)} W^{(l)}) \\ H_2^{(l+1)} &= \sigma(\hat{D}^{(-1/2)} \hat{A} \hat{D}^{(-1/2)} H_2^{(l)} W^{(l)}) \\ H_3^{(l+1)} &= \sigma(\hat{D}^{(-1/2)} \hat{A} \hat{D}^{(-1/2)} H_3^{(l)} W^{(l)}) \end{aligned} \quad (9)$$

where $H^{(l)}$ denotes the feature matrix of the l th layer, $W^{(l)}$ represents the weights of the l th layer, and σ signifies the activation function, with ReLU specifically employed. The symmetric Laplacian matrix operates on the feature matrix of the l th layer along with the weights, computing the feature matrix of the $l + 1$ th layer through the activation function.

3.8. Weather variable

After the propagation process, the three outputs are two-dimensional matrices with the same shape as the feature matrices. These are extended in one-dimensional shape, interconnected, and additionally linked with meteorological variables.

$$\hat{y} = \sigma(W_1 \text{vec}(H_1) + W_2 \text{vec}(H_2) + W_3 \text{vec}(H_3) + W_4 x) \quad (10)$$

Here, \hat{y} denotes the predicted bicycle demand. $\text{vec}(H_1)$ refers to the output of the GCN model that integrates the spatial matrix, stretched into a one-dimensional shape. Meanwhile, $\text{vec}(H_2)$

and $\text{vec}(H_3)$ represent the outputs of the GCN model encompassing the daily and weekly historical rental matrices, respectively, also stretched into one-dimensional shapes. x represents weather variables such as ‘precipitation,’ ‘maximum temperature,’ ‘average temperature,’ and others. W_1, W_2, W_3, W_4, W_5 symbolize the weights of the model. The evaluation criterion employed in this study was the Root Mean Square Error (RMSE) loss function.

3.9. Hyperparameters tuning

In this study, various hyperparameters were adjusted to enhance and optimize the performance of the developed GCN model. Various hyperparameters were considered, including the GCN hidden layer, FCN hidden layer, learning rate, daily temporal window size, weekly temporal window size, epochs, weight decay, and others.

The GCN model developed in this study was implemented using PyTorch and Python. Hyperparameter tuning involved researchers specifying all parameter combinations three times to identify the combination that resulted in the lowest average RMSE value for the test set. To reduce the number of parameter combinations, several measures were taken. The GCN’s hidden layer count included parameters of up to 3 (Ling, 2021), while the FCN hidden layer count was manually set to 2 after identifying no significant difference when the hidden layer count exceeded 2. The sizes for the historical daily and weekly rental windows were chosen based on criteria outlined in previous literature (Kim et al., 2022), selecting daily rental window sizes from 3 to 5 days and weekly rental window sizes from 1 to 5 weeks. Both the GCN and FCN models employed the rectified linear unit (ReLU) as the activation function, and Adam Optimizer was utilized as the optimizer (Table 3).

Hyperparameter	Value
GCN hidden layer	2
FCN hidden layer	2
Day	3
Week	4
Epoch	300
Learning rate	0.001

Table 3. Description of hyperparameter values

Furthermore, hyperparameter tuning for weather variables was undertaken. Following the precedent literature, the precipitation variable was initially included (L. Chen et al., 2016; Frade & Ribeiro, 2014; Hulot et al., 2018; Kim et al., 2022; Thomas, Jaarsma, & Tutert, 2009; C. Xu et al., 2018), and subsequently, temperature variables were definitively incorporated. Predictive performance was then compared based on the presence or absence of other variables. The outcomes revealed that the introduction of variables beyond precipitation (mm) and average temperature (°C) either yielded similar predictive performance or, in some cases, exhibited lower performance. Consequently, a decision was made to include only the precipitation (mm) and average temperature (°C) variables in our model.

4. Results

4.1. Comparative analysis

A comparative analysis was undertaken to evaluate the performance of the GCN-pro model relative to other models, focusing on differences in performance when utilizing various types

of data for demand prediction. The evaluation included models that process time series data, such as LSTM, CNN, SARIMAX, and lightGBM, in comparison to our GCN-pro model, which incorporates spatial data. The analysis also considered the performance of the GCN-UP model, which utilizes graph data as indicated in previous research (Kim et al., 2022). The GCN-pro model demonstrated superior performance across all comparisons. Excluding the GCN-pro, the performance rankings of the other models were, in descending order of effectiveness, GCN-UP, CNN, LSTM, lightGBM, and ARIMA (Table 4). This paragraph highlights the exceptional performance of the GCN-pro model in the context of demand prediction, based on the comparative results.

Model	GCN-pro	GCN-UP	LSTM	CNN	lightGBM	ARIMA
RMSE	12.25	15.71	24.24	16.9	25.1	29.8

Table 4. Performance comparison of different models

4.2. GCN-pro vs CNN

Figure 3 presents a comparison between the GCN-pro and CNN models, focusing on their ability to predict rental quantities influenced by various spatiotemporal features, which exhibit significant daily fluctuations. Accurately capturing these fluctuations and predicting actual rental quantities is essential for demand prediction models. In this study, a detailed comparison of actual versus predicted values for both models was conducted across the September test set. Up until September 20th, both the GCN-pro and CNN models showed commendable accuracy in forecasting rental quantities. However, they faced challenges in mirroring the actual rental quantities beyond this date. Prior to September 20th, the GCN-pro model consistently outperformed the CNN model in capturing daily trends and forecasting demand. Yet, during the period from September 9th to 12th, the performance of the GCN-pro model dipped, which is attributed to the exceptional circumstances surrounding the Chuseok holiday period. This led to significant deviations in bicycle usage patterns compared to regular days. Despite these challenges, the GCN-pro model demonstrated a better alignment with actual values compared to the CNN model, underscoring its superior predictive capability in the context of demand prediction.

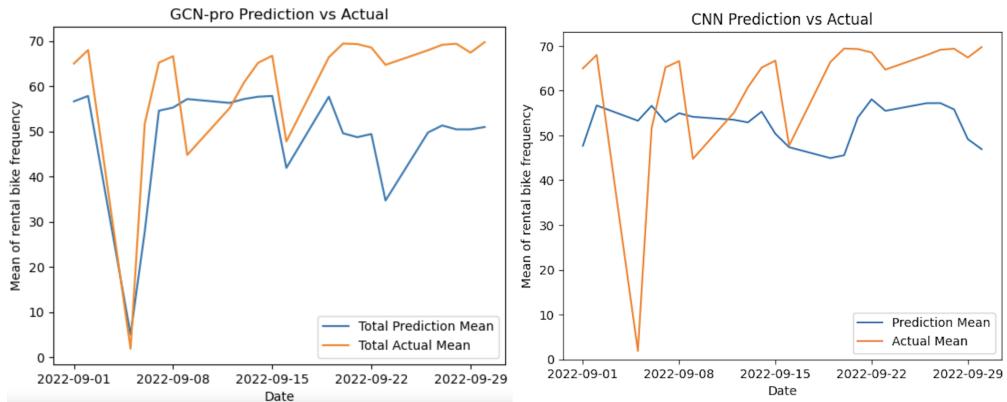


Figure 3. Comparison of prediction performance: CNN vs. GCN-pro

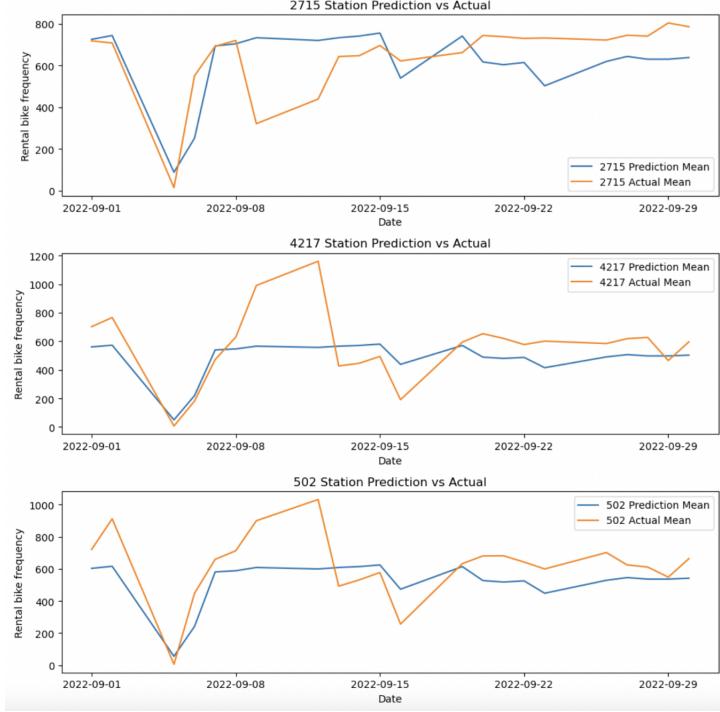


Figure 4. Predicted graphs for the top 3 stations in rental volume using GCN-pro

4.3. Prediction in Station-level

At the individual station level, predictive performance of the GCN-pro model in forecasting bicycle demand is of paramount importance. The GCN-pro model customizes its forecasts to cater to the demand specifics of each station, encompassing a total of 2,709 stations. Figure 4 depicts the prediction outcomes for the three stations experiencing the highest bicycle usage, where the GCN-pro model showcases a strong capability in accurately tracking and predicting sudden shifts in actual rental trends at these individual stations. Nevertheless, it is important to note that during the period from September 9th to 12th, which coincides with the Chuseok holiday, the model faces difficulties in precisely forecasting the actual demand.

Figure 5 illustrates the RMSE values across various stations in Seoul city, using a color-coded system where red signifies higher RMSEs and blue denotes lower ones. Notably, one of the stations with the highest RMSE is the ‘Shinhan Financial Investment Bicycle Rental Station’. This location is distinguished by a complex set of factors impacting bicycle rental volumes. Directly opposite this station is Yeouido Park and the surrounding area includes residential zones, a bus stop, and the Yeouido Transfer Center. Furthermore, its proximity to the Han River and major bridges marks it as a high-traffic zone. These diverse factors can make accurate demand prediction challenging.

5. Discussion

This study introduces a novel approach to predicting bicycle demand by incorporating station-specific spatial characteristics, an aspect overlooked in previous research. By leveraging a combination of GCN and FCN, this research enhances the predictive accuracy by considering both spatial and temporal data, including rental volumes, weather conditions, and proximate spatial features within approximately 500 meters of each station. The nodes in the GCN-pro model represent individual docking stations, enabling precise daily demand forecasts at the station level. A critical analysis comparing the GCN-pro model’s effectiveness with that of

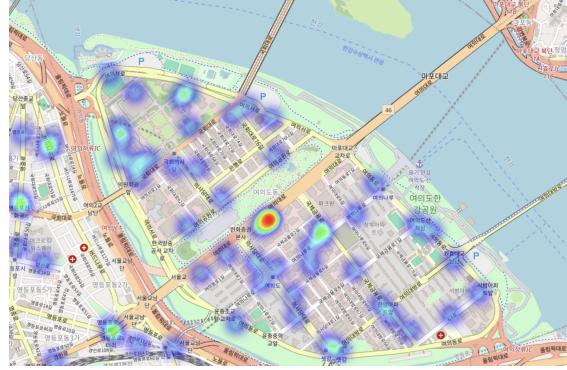


Figure 5. Seoul Station RMSE Heat-map: Red for High, Green-Blue for Low Errors



Figure 6. Map of Bicycle Rental Stations in Areas with High Pedestrian Traffic

* The red dots marked with asterisks represent the rental stations.

existing GCN models, as well as traditional models like LSTM, CNN, lightGBM, and ARIMA, highlights its superior performance. Furthermore, the significance of spatial features on rental demand was quantitatively assessed using the Kruskal-Wallis test, identifying 11 out of 15 spatial features as significant factors. The findings underscore the importance of integrating spatial details at the node level into the graph structure, which proves more effective than models relying solely on temporal dynamics or spatial characteristics included in the adjacency matrix. This refined approach not only advances our understanding of demand prediction but also demonstrates the value of incorporating comprehensive spatial-temporal data for more accurate demand forecasting in public bicycle systems.

While our proposed model exhibits high accuracy, it faces greater challenges in stations situated in areas of significant pedestrian activity, indicative of high foot traffic, similar to challenges observed with other models. For example, as shown in the left panel of Figure A1, Station 829, located at the North End Crossroad near Hangangdaegyo Bridge and Gangbyeonbuk-ro, is surrounded by residential areas including Hangang Daewoo and Daelim Apartment complexes. It is also near parks such as Ichon Hangang Park, Yongsan Children's Park, and Seobinggo Neighborhood Park, with Yongsan and Ichon being the nearest subway stations. Furthermore, the right panel of Figure 7 highlights the rental station in front of Hongik University Station Exit 2, situated in an area experiencing the highest foot traffic among the 20 to 40 age group, who are the main users of the public bike-sharing service. The intense activity level of this demographic makes predictions particularly challenging. These observations highlight the complex and multifaceted challenges in predicting bicycle rental volumes, especially in areas with high variability and significant pedestrian or vehicular traffic, underscoring the limitations of our model in handling such dynamically varied environments.

Acknowledgement

Special thanks to Jisu Shin, the designer, for providing graphical expertise in creating the model summary diagram.

Disclosure statement

An unnumbered section, e.g. \section*{Disclosure statement}, may be used to declare any potential conflict of interest and included *in the non-anonymous version* before any Notes or References, after any Acknowledgements and before any Funding information.

Funding

This research was supported by the Ministry of Education of the Republic of Korea and the NRF (No. 2023S1A5A2A21086671).

References

- Agarwal, A., Ziemke, D., & Nagel, K. (2020). Bicycle superhighway: An environmentally sustainable policy for urban transport. *Transportation Research Part A: Policy and Practice*, 137(January 2019), 519–540. Retrieved from <https://doi.org/10.1016/j.tra.2019.06.015>
- Boufidis, N., Nikiforidis, A., Chrysostomou, K., & Aifadopoulou, G. (2020). Development of a station-level demand prediction and visualization tool to support bike-sharing systems' operators. *Transportation Research Procedia*, 47, 51–58. Retrieved 2024-02-15, from <https://linkinghub.elsevier.com/retrieve/pii/S2352146520302593>
- Cagliero, L., Cerquitelli, T., Chiusano, S., Garza, P., & Xiao, X. (2017). Predicting critical conditions in bicycle sharing systems. *Computing*, 99(1), 39–57.
- Chai, D., Wang, L., & Yang, Q. (2018). Bike flow prediction with multi-graph convolutional networks. *GIS: Proceedings of the ACM International Symposium on Advances in Geographic Information Systems*, 397–400.
- Chen, L., Zhang, D., Wang, L., Yang, D., Ma, X., Li, S., ... Jakubowicz, J. (2016). Dynamic cluster-based over-demand prediction in bike sharing systems. *UbiComp 2016 - Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 841–852.
- Chen, P., Hsieh, H., Su, K., Sigalingging, X. K., Chen, Y., & Leu, J. (2020, June). Predicting station level demand in a bike-sharing system using recurrent neural networks. *IET Intelligent Transport Systems*, 14(6), 554–561. Retrieved 2024-02-15, from <https://onlinelibrary.wiley.com/doi/10.1049/iet-its.2019.0007>
- Cui, X. (2018). Influencing factors of public participation willingness in shared bicycles and intervention strategies. *Journal of Discrete Mathematical Sciences and Cryptography*, 21(6), 1437–1442.
- Du, Y., Deng, F., & Liao, F. (2019, June). A model framework for discovering the spatio-temporal usage patterns of public free-floating bike-sharing system. *Transportation Research Part C: Emerging Technologies*, 103, 39–55. Retrieved 2024-02-15, from <https://linkinghub.elsevier.com/retrieve/pii/S0968090X18306673>
- Fanaee-T, H., & Gama, J. (2014, June). Event labeling combining ensemble detectors and background knowledge. *Progress in Artificial Intelligence*, 2(2-3), 113–127. Retrieved 2024-02-15, from <http://link.springer.com/10.1007/s13748-013-0040-3>
- Feng, S., Chen, H., Du, C., Li, J., & Jing, N. (2018). A hierarchical demand prediction method with station clustering for bike sharing system. *Proceedings - 2018 IEEE 3rd International Conference on Data Science in Cyberspace, DSC 2018*, 829–836.
- Frade, I., & Ribeiro, A. (2014). Bicycle Sharing Systems Demand. *Procedia - Social and Behavioral Sciences*, 111(1999), 518–527. Retrieved from <http://dx.doi.org/10.1016/j.sbspro.2014.01.085>
- Gebhart, K., & Noland, R. B. (2014). The impact of weather conditions on bikeshare trips in Washington, DC. *Transportation*, 41(6), 1205–1225.

- Giot, R., & Cherrier, R. (2015). Predicting bikeshare system usage up to one day ahead. *IEEE SSCI 2014: 2014 IEEE Symposium Series on Computational Intelligence - CIVTS 2014: 2014 IEEE Symposium on Computational Intelligence in Vehicles and Transportation Systems, Proceedings*, 22–29.
- Guo, R., Jiang, Z., Huang, J., Tao, J., Wang, C., Li, J., & Chen, L. (2019). BikeNet: Accurate bike demand prediction using graph neural networks for station rebalancing. *Proceedings - 2019 IEEE SmartWorld, Ubiquitous Intelligence and Computing, Advanced and Trusted Computing, Scalable Computing and Communications, Internet of People and Smart City Innovation, Smart-World/UIC/ATC/SCALCOM/IOP/SCI 2019*, 686–693.
- Guo, S., Lin, Y., Feng, N., Song, C., & Wan, H. (2019). Attention based spatial-temporal graph convolutional networks for traffic flow forecasting. *33rd AAAI Conference on Artificial Intelligence, AAAI 2019, 31st Innovative Applications of Artificial Intelligence Conference, IAAI 2019 and the 9th AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2019*, 922–929.
- Hammond, D. K., Vandergheynst, P., & Gribonval, R. (2011). Wavelets on graphs via spectral graph theory. *Applied and Computational Harmonic Analysis*, 30(2), 129–150. Retrieved from <http://dx.doi.org/10.1016/j.acha.2010.04.005>
- Hulot, P., Aloise, D., & Jena, S. D. (2018). Towards station-level demand prediction for effective rebalancing in bike-sharing systems. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 378–386.
- Kim, T. S., Lee, W. K., & Sohn, S. Y. (2022). Erratum: Correction: Graph convolutional network approach applied to predict hourly bike-sharing demands considering spatial, temporal, and global effects (PloS one (2019) 14 9 (e0220782)). *PloS one*, 17(3), e0266221.
- Kipf, T. N., & Welling, M. (2017). Semi-supervised classification with graph convolutional networks. *5th International Conference on Learning Representations, ICLR 2017 - Conference Track Proceedings*, 1–14.
- Kloimüllner, C., Papazek, P., Hu, B., & Raidl, G. R. (2014). Balancing bicycle sharing systems: An approach for the dynamic case. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 8600, 73–84.
- Li, Q., Fuerst, F., & Luca, D. (2023). Do shared E-bikes reduce urban carbon emissions? *Journal of Transport Geography*, 112(September), 103697. Retrieved from <https://doi.org/10.1016/j.jtrangeo.2023.103697>
- Lin, L., He, Z., & Peeta, S. (2018). Predicting station-level hourly demand in a large-scale bike-sharing network: A graph convolutional neural network approach. *Transportation Research Part C: Emerging Technologies*, 97(November), 258–276. Retrieved from <https://doi.org/10.1016/j.trc.2018.10.011>
- Lin, P., Weng, J., Hu, S., Alivanistos, D., Li, X., & Yin, B. (2020). Revealing spatio-temporal patterns and influencing factors of dockless bike sharing demand. *IEEE Access*, 8, 66139–66149. Retrieved 2024-02-15, from <https://ieeexplore.ieee.org/document/9054982/>
- Ling, N. (2021). Improved Hyperparameter Tuning for Graph Learning with Warm-Start Configuration
Improved Hyperparameter Tuning for Graph Learning with Warm-Start Configuration.
- Liu, J., Sun, L., Li, Q., Ming, J., Liu, Y., & Xiong, H. (2017). Functional zone based hierarchical demand prediction for bike system expansion. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Part F1296*, 957–966.
- Park, J., Honda, Y., Fujii, S., & Kim, S. E. (2022, March). Air pollution and public bike-sharing system ridership in the context of sustainable development goals. *Sustainability*, 14(7), 3861. Retrieved 2024-02-15, from <https://www.mdpi.com/2071-1050/14/7/3861>
- Ramesh, A. A., Nagisetti, S. P., Sridhar, N., Avery, K., & Bein, D. (2021). Station-level Demand Prediction for Bike-Sharing System. *2021 IEEE 11th Annual Computing and Communication Workshop and Conference, CCWC 2021*, 916–921.
- Salaken, S. M., Hosen, M. A., Khosravi, A., & Nahavandi, S. (2015). Forecasting bike sharing demand using fuzzy inference mechanism. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 9491, 567–574.
- Singhvi, D., Singhvi, S., Frazier, P. I., Henderson, S. G., O'Mahony, E., Shmoys, D. B., & Woodard, D. B. (2015). Predicting bike usage for New York city's bike sharing system. *AAAI Workshop - Technical Report, WS-15-06*, 110–114.
- Thomas, T., Jaarsma, R., & Tutert, B. (2009). Temporal variations of bicycle demand in the Netherlands : The influence of weather on cycling. *Transportation Research Board 88th Annual Meeting*, 1–17.
- V E, S., & Cho, Y. (2020). A rule-based model for Seoul Bike sharing demand prediction using weather data. *European Journal of Remote Sensing*, 53(sup1), 166–183. Retrieved from

- <https://doi.org/10.1080/22797254.2020.1725789>
- Xia, X., Jiang, H., & Wang, J. (2022). Analysis of user satisfaction of shared bicycles based on SEM. *Journal of Ambient Intelligence and Humanized Computing*, 13(3), 1587–1601. Retrieved from <https://doi.org/10.1007/s12652-019-01422-y>
- Xu, C., Ji, J., & Liu, P. (2018). The station-free sharing bike demand forecasting with a deep learning approach and large-scale datasets. *Transportation Research Part C: Emerging Technologies*, 95(July), 47–60. Retrieved from <https://doi.org/10.1016/j.trc.2018.07.013>
- Xu, H., Duan, F., & Pu, P. (2019). Dynamic bicycle scheduling problem based on short-term demand prediction. *Applied Intelligence*, 49(5), 1968–1981.
- Xu, H., Ying, J., Wu, H., & Lin, F. (2013, March). Public bicycle traffic flow prediction based on a hybrid model. *Applied Mathematics & Information Sciences*, 7(2), 667–674. Retrieved 2024-02-15, from <http://www.naturalspublishing.com/Article.asp?ArtcID=2335>
- Xu, T., Han, G., Qi, X., Du, J., Lin, C., & Shu, L. (2020). A Hybrid Machine Learning Model for Demand Prediction of Edge-Computing-Based Bike-Sharing System Using Internet of Things. *IEEE Internet of Things Journal*, 7(8), 7345–7356.
- Yang, L., Fei, S., Jia, H., Qi, J., Wang, L., & Hu, X. (2023, September). Study on the relationship between the spatial distribution of shared bicycle travel demand and urban built environment. *Sustainability*, 15(18), 13576. Retrieved 2024-02-15, from <https://www.mdpi.com/2071-1050/15/18/13576>
- Yang, X., & He, S. (2020, August). *Towards dynamic urban bike usage prediction for station network reconfiguration*. arXiv. Retrieved 2024-02-15, from <http://arxiv.org/abs/2008.07318> [arXiv:2008.07318 [cs, stat]]
- Yu, B., Yin, H., & Zhu, Z. (2018). Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting. *IJCAI International Joint Conference on Artificial Intelligence*, 2018-July, 3634–3640.
- Zhang, H., Zhuge, C., Jia, J., Shi, B., & Wang, W. (2021, September). Green travel mobility of dockless bike-sharing based on trip data in big cities: A spatial network analysis. *Journal of Cleaner Production*, 313, 127930. Retrieved 2024-02-15, from <https://linkinghub.elsevier.com/retrieve/pii/S095965262102148X>
- Zhang, S., Tong, H., Xu, J., & Maciejewski, R. (2019). Graph convolutional networks: a comprehensive review. *Computational Social Networks*, 6(1). Retrieved from <https://doi.org/10.1186/s40649-019-0069-y>
- Zhao, S., Zhao, K., Xia, Y., & Jia, W. (2022). Hyper-clustering enhanced spatio-temporal deep learning for traffic and demand prediction in bike-sharing systems. *Information Sciences*, 612, 626–637. Retrieved from <https://doi.org/10.1016/j.ins.2022.07.054>
- Zhou, Q., Wong, C. J., & Su, X. (2020). Machine learning approach to quantity management for long-term sustainable development of dockless public bike: Case of Shenzhen in China. *Journal of Advanced Transportation*, 2020(2).
- Zhou, Y., Wang, L., Zhong, R., & Tan, Y. (2018). A markov chain based demand prediction model for stations in bike sharing systems. *Mathematical Problems in Engineering*, 2018, 1–8. Retrieved 2024-02-15, from <https://www.hindawi.com/journals/mpe/2018/8028714/>
- Zi, W., Xiong, W., Chen, H., & Chen, L. (2021). TAGCN: Station-level demand prediction for bike-sharing system via a temporal attention graph convolution network. *Information Sciences*, 561(109), 274–285. Retrieved from <https://doi.org/10.1016/j.ins.2021.01.065>

6. Appendices

Appendix A. Histogram of RMSE on Comaparative analysis

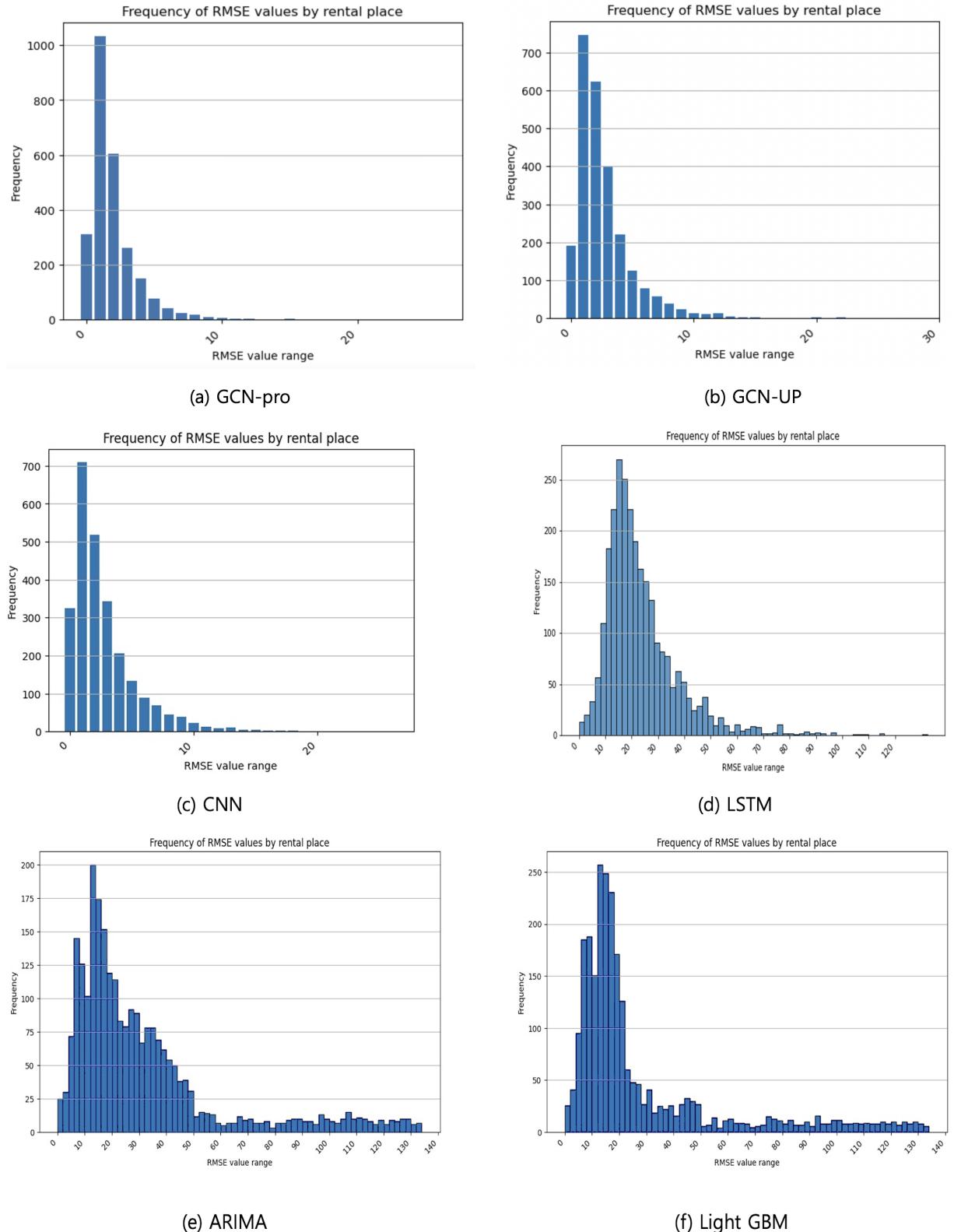


Figure A1. This is a histogram depicting the Root Mean Square Error (RMSE) values of daily predictions for bicycle rental stations. The horizontal axis represents the range of RMSE values, while the vertical axis indicates the frequency of stations falling within each range. When predicting with the GCN-pro model, it is evident that the majority of stations have RMSE values below 10. In contrast, models such as LSTM, ARIMA, and lightGBM exhibit a substantial presence of stations with significant discrepancies between predicted and actual values.