

Multi-Branch Neural Networks for Predicting Shared Bikes

Huigyu Yang¹, Syed M. Raza², Duc-Tai Le², Dongsoo S. Kim³, Hyunseung Choo²

¹Dept. of Superintelligence Engineering, *Sungkyunkwan University, Suwon, South Korea*

²Dept. of Electrical and Computer Engineering, *Sungkyunkwan University, Suwon, South Korea*

³Dept. of Electrical and Computer Engineering, *Indiana University-Purdue University, Indianapolis, USA*
{huigyu, syed.m.raza, ldtai, choo}@skku.edu, dskim@iupui.edu

Abstract—Bike sharing systems are one of the sharing economies that aims to relieve the limited parking space of cities and to elevate public transportation. The sharing systems provide efficient user mobility as users determine the rental period of a sharing service according to their demand. They distinguish the service implementations into dock-based and dockless types. A dock refers a designated place and facilities where users rent or return the bikes. The dock requires the service provider to maintain its facilities such as the number of available bikes at each station (i.e., a group of docks). The dockless service allows user to start and end the rental service from anywhere. facilities such as the number of available bikes at each station (i.e., a group of docks). The dockless service allows user to start and end the rental service from anywhere.

Index Terms—Bike Sharing, Time-series Prediction, Multi-Branch Neural Network

I. INTRODUCTION

Bike sharing systems are one of the sharing economies that aims to relieve the limited parking space of cities and to elevate public transportation. The sharing systems provide efficient user mobility as users determine the rental period of a sharing service according to their demand. They distinguish the service implementations into dock-based and dockless types. A dock refers a designated place and facilities where users rent or return the bikes. The dock requires the service provider to maintain its facilities such as the number of available bikes at each station (i.e., a group of docks). The dockless service allows user to start and end the rental service from anywhere. The users put additional efforts to find the available bikes, and the maintenance of the dockless service requires higher cost due to similar reason.

In most cities, residential or commercial areas are separated, so user movements are concentrated in specific areas depending on the region and time. Dock-based systems rent or return bikes at fixed service locations, resulting in insufficient or overflowing demand-supply balance problems depending on time and place. To address this problem, the system operator redeploys the bikes at regular intervals to distribute evenly across the stations. Demand-supply imbalances occur nonlinearly in terms of time, and time-consistent relocation causes overhead in service delivery. The prediction of the

distribution status of a shared bicycle can estimate the time of need for relocation and enable efficient operating costs.

Early studies on bike sharing system have correlated user demands with environmental factors such as weather, public transportation, temporal factors, and safety [3, 4]. The models conduct convolutional neural networks to learn spatial characteristics of target region data and recurrent neural network families to utilize temporal characteristics of mobility [5]. These models have evolved to improve prediction accuracy through graph neural networks [6] or clustering techniques [7] which extract the correlation of the docks or various influential factors. These models have evolved to improve prediction accuracy through graph neural networks [6] or clustering techniques [7] which extract the correlation of the docks or various influential factors.

In recent studies, the prediction models improved accuracy performance using fusion layers and residual connections on multiple input data such as geographical or meteorological information, and mobility patterns [5][8]. The limitation of these studies is that prediction results or performances of deep learning models are mainly presented rather than focused on rebalancing problem which remains as a challenging task. In recent studies, the prediction models improved accuracy performance using fusion layers and residual connections on multiple input data such as geographical or meteorological information, and mobility patterns [5][8]. The limitation of these studies is that prediction results or performances of deep learning models are mainly presented rather than focused on rebalancing problem which remains as a challenging task.

Our proposed model predicts In most cities, residential or commercial areas are separated, so user movements are concentrated in specific areas depending on the region and time. Dock-based systems rent or return bikes at fixed service locations, resulting in insufficient or overflowing demand-supply balance problems depending on time and place. To address this problem, the system operator redeploys the bikes at regular intervals to distribute evenly across the stations. Demand-supply imbalances occur nonlinearly in terms of time, and time-consistent relocation causes overhead in service delivery. The prediction of the distribution status of a shared bicycle can estimate the time of need for relocation and enable efficient operating costs.

II. PREDICTION MODEL FOR SHARED BIKES

A. Bike Sharing Dataset

Bike sharing services are actively used in various countries and cities such as Chicago, Washington, New York, Singapore, and Taipei. The system explicitly collects data generated from users' service use and provides anonymized data as a public dataset for research purposes in a standardized format. Datasets commonly include datetime, coordinates of rental and return locations, usage time, and some datasets additionally provide metadata such as season, weather, temperature, and humidity. Chicago's bike-sharing system also provides logs measuring the number of bicycles in 611 stations every 10 minutes and data recording anonymized user movements in coordinates. Due to its rich data properties, we leverage datasets collected in Chicago for training and testing prediction models.

Bike sharing services are actively used in various countries and cities such as Chicago, Washington, New York, Singapore, and Taipei. The system explicitly collects data generated from users' service use and provides anonymized data as a public dataset for research purposes in a standardized format. Datasets commonly include datetime, coordinates of rental and return locations, usage time, and some datasets additionally provide metadata such as season, weather, temperature, and humidity. Chicago's bike-sharing system also provides logs measuring the number of bicycles in 611 stations every 10 minutes and data recording anonymized user movements in coordinates. Due to its rich data properties, we leverage datasets collected in Chicago for training and testing prediction models.

B. Multi-branch Neural Networks

The proposed model predicts the available bikes $\hat{s}(t+1)$ and classifies the events $e(t+1)$ of the stations. Two LSTM layers of the model extract temporal features from dynamic patterns of the input sequence. The extracted features are represented in a 256-dimensional latent vector, where the dimensionality is determined by the LSTM cell. To result various types of outputs from the model, the latent vector is duplicated and passed to two different Fully-Connected (FC) layers. As shown in Fig. 1, FC Layers with Cross-Entropy (CE) and Root Mean Square Error (RMSE) convert the latent vector into the probability vector of events and the numbers of bikes, respectively. In the training phase, calculated losses from the event label and the bike amount individually backpropagate through each FC layer. This backpropagation jointly updates trainable parameters in LSTM layers after two FC layers.

The proposed model predicts the available bikes $\hat{s}(t+1)$ and classifies the events $e(t+1)$ of the stations. Two LSTM layers of the model extract temporal features from dynamic patterns of the input sequence. The extracted features are represented in a 256-dimensional latent vector, where the dimensionality is determined by the LSTM cell. To result various types of outputs from the model, the latent vector is duplicated and passed to two different Fully-Connected (FC) layers. As

shown in Fig. 1, FC Layers with Cross-Entropy (CE) and Root Mean Square Error (RMSE) convert the latent vector into the probability vector of events and the numbers of bikes, respectively. In the training phase, calculated losses from the event label and the bike amount individually backpropagate through each FC layer. This backpropagation jointly updates trainable parameters in LSTM layers after two FC layers.

The proposed model predicts the available bikes $\hat{s}(t+1)$ and classifies the events $e(t+1)$ of the stations. Two LSTM layers of the model extract temporal features from dynamic patterns of the input sequence. The extracted features are represented in a 256-dimensional latent vector, where the dimensionality is determined by the LSTM cell. To result various types of outputs from the model, the latent vector is duplicated and passed to two different Fully-Connected (FC) layers. As shown in Fig. 1, FC Layers with Cross-Entropy (CE) and Root Mean Square Error (RMSE) convert the latent vector into the probability vector of events and the numbers of bikes, respectively. In the training phase, calculated losses from the event label and the bike amount individually backpropagate through each FC layer. This backpropagation jointly updates trainable parameters in LSTM layers after two FC layers.

III. RESULTS AND ANALYSES

Before you begin to format your paper, first write and save the content as a separate text file. Complete all content and organizational editing before proofreading, spelling and grammar.

Keep your text and graphic files separate until after the text has been formatted and styled. Do not number text heads— \LaTeX will do that for you.

IV. CONCLUSION AND FUTURE WORK

Fig. 1. Example of a figure caption.

ACKNOWLEDGMENT

This research was supported by the MSIT(Ministry of Science, ICT), Korea, under the ICT Creative Consilience program (IITP-2021-2020-0-01821) and High-Potential Individuals Global Training Program (IITP-2021-0-02132) supervised by the IITP(Institute for Information & Communications Technology Planning & Evaluation)

REFERENCES

- [1] G. Eason, B. Noble, and I. N. Sneddon, "On certain integrals of Lipschitz-Hankel type involving products of Bessel functions," *Phil. Trans. Roy. Soc. London*, vol. A247, pp. 529–551, April 1955.
- [2] J. Clerk Maxwell, *A Treatise on Electricity and Magnetism*, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.
- [3] I. S. Jacobs and C. P. Bean, "Fine particles, thin films and exchange anisotropy," in *Magnetism*, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.
- [4] K. Elissa, "Title of paper if known," unpublished.
- [5] R. Nicole, "Title of paper with only first word capitalized," *J. Name Stand. Abbrev.*, in press.

- [6] Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, "Electron spectroscopy studies on magneto-optical media and plastic substrate interface," IEEE Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetism Japan, p. 301, 1982].
- [7] M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.

IEEE conference templates contain guidance text for composing and formatting conference papers. Please ensure that all template text is removed from your conference paper prior to submission to the conference. Failure to remove the template text from your paper may result in your paper not being published.