Bicycle Sharing Systems meet AI: forecasting bicycles availability and decision-making.

Jože M. Rožanec, Tadej Krivec
Jožef Stefan International Postgraduate School
Jamova cesta 39, 1000 Ljubljana, Slovenia
{joze.rozanec, tadej.krivec}@ijs.si

Vid Keršič
University of Maribor
Faculty of Electrical Engineering and Computer Science
UM FERI, Koroška cesta 46, 2000 Maribor, Slovenia
vid.kersic@um.si

Larsen Cundrič
Longevize BV
l.cundric@gmail.com

Blaž Stojanovič
blaz.sto@gmail.com

Marko Zeman, Ivan Bratko
University of Ljubljana
Faculty of Computer and Information Science
Večna pot 113, 1000 Ljubljana, Slovenia
{marko.zeman, bratko}@fri.uni-lj.si

Abstract. *With the ubiquitous increase in the number of people in cities, there is a growing need for sustainable transport possibilities. Smart cities should provide environment-friendly ways to travel inside the city. One of the most nature-preserving ways to travel is using bicycles, which is often encouraged by public bicycle sharing systems, which are present in many cities around the globe. While these systems are usually easily accessible, they still lack optimization regarding bicycle availability across stations. This paper contributes towards the optimization of the bicycle sharing system in Ljubljana, Slovenia. We developed classification and regression machine learning models for predicting the emptiness and occupancy across bicycle stations in near future. These predictions allow for the caretakers of the system to intervene on time and provide enough bicycles across all stations.*

Keywords. Smart city, sustainable transport, bicycle sharing, optimization

1 Introduction

The world is facing an environmental crisis, where the ecosystems are changing rapidly. In 2018, the intergovernmental body reported that we are facing irreversible consequences due to the global greenhouse gas (GHG) emissions (Masson-Delmotte et al. (2018)). For that reason, the goal is to address climate change by preparing for unavoidable consequences or reducing emissions. One of the areas where the benefits of mitigating the GHG emissions would be significant is transport.

Transportation systems account for a quarter of energy-related CO_2 emissions (Masson-Delmotte et al. (2018)). One of the strategies to reduce GHG emissions in transport is a modal shift to lower-carbon options (Rolnick et al. (2022)). One of the best ways to substitute the carbon emission intensive modes of transport in densely populated areas is with bicycles.

The first bicycle-sharing system (BSS) was introduced in 1965 in Amsterdam with the *White Bike Plan* bicycle-sharing program. Since then, Shaheen et al. (2010) reported that more than 130.000 bicycles have been made available through such programs worldwide. In all such programs, the users can borrow a bicycle from one of the stations and return it to some station within a specific time frame.

Among the biggest challenges of the BSS is the (re)balancing of the bicycles at the parking stations (Shaheen et al. (2010)). For example, in the morning and afternoon hours, the direction of travel by bicycle is usually heavily biased in one direction. This bias, in turn, causes an imbalance at the stations, where some stations are full and others are empty. The employed workers then redistribute bicycles from occupied stations to less occupied ones. However, the open question remains of how and when the employees should take action. Machine learning (ML) provides means to solve such a problem.

ML algorithms have already achieved great success in various forecasting tasks. There is a big motivation to use ML for mitigating climate change. For example, ML is used to forecast electricity demand, photovoltaic production, extreme events, and behavior change (Rolnick et al. (2022)). For the task at hand, ML algorithms can assist decision-making in the bicycle redistribution problem.

In this paper, we present an approach to classifying decision-making options (adding/removing bicycles from stations) and an approach to forecasting the number of bicycles at the parking stations. We do so considering multiple forecasting horizons. Our contributions involve developing and comparing classification models for decision-making options and regression models for bicycle availability, which can be used in real-time by the BicikeLJ BSS to assist the rebalancing process.

The rest of this paper is structured as follows: Section 2 presents related work, and Section 3 describes the use-case of a BSS used in Ljubljana, Slovenia. Section 4 introduces the methodology we followed to analyze the data, create the dataset, the corresponding features, and how we trained and evaluated the forecasting models. Section 5 presents the results, while Section 6 offers our conclusions and provides an outline for future work.

2 Related Work

Shaheen et al. (2010) provide an overview of BSS and related problems and experiences. Bicycle imbalance in a network of bicycle stations is a critical problem. Therefore bicycle redistribution is a necessary functionality of such systems. Usually, it is done by trucks or other vehicles to move bicycles to high-demand stations. Efficient maintenance of bicycle balance requires online information support, automated prediction of the number of bicycles and demand for a bicycle at individual stations, and planning optimal interventions to restore the balance.

Many works (Yang et al. (2016); Almannaa et al. (2020); Ashqar et al. (2017); Cagliero et al. (2017); Ashqar et al. (2021)) report on experiments with the automated prediction of the number of bicycles at different stations, using various statistical and ML techniques for learning to predict from past BSS data. Yang et al. (2016) proposed a prediction method that consists of (1) predicting the check-in numbers at individual stations, given the current data about check-out bicycles and their locations, taking into account the probabilities of bicycles currently traveling between pairs of stations and travel times; and (2) estimating the probability distribution of checking out a bicycle at given station depending on the time of day, kind of day, type of station location, weather conditions, among other factors. Using historical data, this is done by the Random Forest ML method (Breiman (2001)). They conducted an experimental evaluation of their approach regarding prediction accuracy using past data observed in the BSS of the Chinese city of Hangzhou, supposedly the world-largest BSS. They compared Random Forest results with three more traditional techniques for time-series prediction and found that Random Forest outperformed these baseline methods.

Almannaa et al. (2020); Ashqar et al. (2017, 2021) report on experiments in the prediction of the number of bicycles at bicycle stations using historical data in the BSS of the San Francisco Bay area. Prediction methods used in Almannaa et al. (2020) were dynamic linear models and Random Forest. Both methods produced similar results in terms of prediction accuracy. For the fifteen-minute predictions horizon of bicycle counts at a station, the average prediction error was about 2% of the station's total bicycle capacity, and the average error for the two-hour prediction horizon was 6%. Cagliero et al. (2017) developed a method for predicting critical conditions in a BSS regarding bicycle counts, integrating the Bayesian learning and Association learning techniques. The method was evaluated on historical bicycle numbers data for the New York City BSS.

The above papers are mainly concerned with predicting bicycle counts but not explicitly with planning optimal bicycle redistribution interventions. Seo et al. (2022) formulated the bicycle redistribution problem as a Markov decision problem. Then the choice of best actions (bicycle moving operations) according to a suitable cost function in a stochastic dynamic environment can be addressed by techniques of reinforcement learning (Sutton and Barto (2018)). The proposed solution suggests the best action every ten minutes, which depends on the observed current state of the BSS system and predicted future demands. Predictions are made by the Random Forest method. This approach was experimentally evaluated on the BSS of Yeouido Island District in South Korea (thirty-one bicycle stations). The experiment used historical data covering about one year from this BSS. The proposed solution to bicycle rebalancing has a clear advantage over previously proposed benchmark strategies in responding to dynamic changes in bicycle demand and thus reducing the chances of unmet demand.

Related work extends to studying various other problems and questions related to BSS. For example, Yoon et al. (2012) described a smartphone application for the Dublin BSS, which recommends the most suitable pair of bicycle stations concerning the user's current needs and the availability of bicycles, and the possibility of bicycle return at different stations. Shang et al. (2021) analyses the impact of the COVID-19 pandemic on the behavior of users in BSS. It is reported that notable effects of the pandemic on users' behavior are mostly environment friendly.

3 Use Case

We performed our research based on a real-world use case based on data from BicikeLJ, a BSS available in Ljubljana, the capital city of Slovenia. BicikeLJ was introduced in 2011. Users pay a small fee to register to the platform and benefit from the BSS. Bicycle usage under an hour is granted for free, while additional charges apply per hour if used for a more extended period. BicikeLJ provides near-real-time information on the number of bicycles available at each station and makes them available to the public through a website[1].

[1]A map with overlayed information regarding the number of bicycles available at each station is provided at the following web page: `https://www.bicikelj.si/en/mapping`

While this is not always the case, the optimal scenario of a BSS would be that there is always (a) a bicycle available for anyone willing to use it and (b) some parking slot at the station so that the bicycle can be parked. Therefore, in this work, we explore two kinds of prediction models: (i) classification models to predict decision-making options (whether we need to add bicycles, remove them, or no action is required), and (ii) regression models to predict the number of bicycles at the parking station. We envision such models could be used to:

- proactively alert decision-makers and ensure all stations have an adequate number of bicycles and parking slots. Such information could be used as inputs to vehicle routing optimization problems if some truck is used to re-balance the bicycles among the stations.
- provide forecasts to the users and ensure they can better understand if some bicycle or parking slot will be available upon arrival at the station. Such forecasts could be complemented with (a) soft reservations to gather demand data and proactively handle mismatches between supply and demand, and (b) some gamification to incentivize people moving bicycles between stations for certain benefit (Wu (2020)).

For the purpose of this research, we considered multiple forecasting horizons. Such forecasts enable us to consider bicycle-sharing dynamics over time and understand how the quality of the forecasts degrades over time.

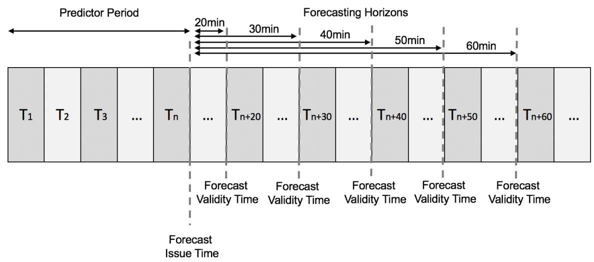

Figure 1: Diagram showing the predictor period taken into account to build the forecasting models, the forecast issue time, and the multiple forecasting horizons considered for this use case. Each model forecasts for a specific horizon and has a particular forecast validity time.

4 Methodology

The subsections below explain the methodology we followed to build and evaluate our forecasting models. We describe how data was collected and analyzed, the features we created, how we defined the target labels and baseline models, and how we trained and evaluated the models.

4.1 Data collection and analysis

The experiment was conducted based on data collected from public APIs provided by the BicikeLJ BSS. Since there are no publicly available historical datasets, we obtained the data from the developer's API and public endpoints of the system. We collected data between March 9th 2022 to May 8th 2022. The final dataset had minor gaps due to networking errors during the web scraping. The data we collected contained two types of records: (i) station states and (ii) bicycle states. Each data record for (i) contains station ID, station name, station address, latitude, longitude, capacity, and the number of free and occupied spaces, while the data from (ii) provide information regarding the bicycle ID, and status, timestamp, and station ID. We collected data for a total of 82 stations and found that there are nearly 800 bicycles offered in the BSS. We then created a single dataset informing the number of bicycles at each station for each reading, which is available at a five-minute frequency. We analyzed the data to find and understand patterns when visualizing the (i) time series (e.g., see Fig. 2), (ii) histograms describing the number of bicycles available at each station (e.g., see Fig. 3), and (iii) heatmaps to understand usual movements of bicycles between stations.

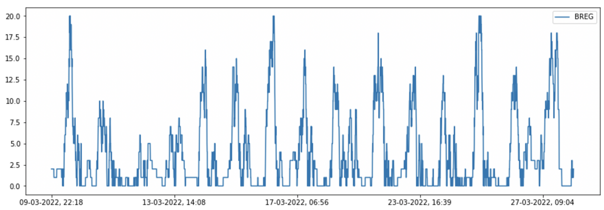

Figure 2: Time series obtained from the bicycle count over the time at the *Breg* station.

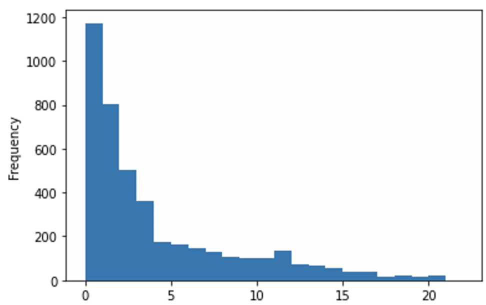

Figure 3: Histogram showing the frequency of bicycle availability for the *Breg* station.

4.2 Feature extraction

While we considered two problems (forecasting decision-making and the number of expected bicycles per station for a given point in time), we considered the same time series features for both. In particular, we consider that the future number of bicycles available at a particular station is closely related to the current demand. Therefore, we computed two types of features (see Table 1): (a) features indicating how close we are to reaching the full stations' bicycle parking capacity (features 1-4) and (b) features indicating whether the number of bicycles in the station is growing or not (features 5-6). The features were created considering relative values to ensure the same features would generalize well between local and global models and make them comparable.

4.3 Creating target labels for classification

One of the possible use cases related to BSS is to assist decision-making. We frame it as a classification problem, aiming to predict whether three possible states will be reached in the future: (a) no bicycles will be available (therefore need to be supplied), (b) no parking slots will be available (therefore bicycles should be removed), or (c) no action is required, given enough bicycles and parking slots will exist. To determine these three labels, we analyzed each station and the preceding number of bicycles before no bicycles (L_{NB}) or parking slots were found (L_{NPS}). Therefore, three classes were assigned: (a) add bicycles if a less or equal amount of bicycles in a given station are found than the reference amount L_{NB}, (b) do not perform any action, and (c) remove bicycles if an equal or greater amount of bicycles in a given station is found than the reference amount L_{NPS}. The dataset has shown a strong class imbalance.

4.4 Baseline models

We created two baseline models, one for the classification and one for the regression problem. We followed a similar procedure for both baselines to the one described in Section 4.3. However, instead of considering state transitions from bicycles and parking slots available to their deficit, we considered the corrective actions that overturned the deficit of bicycles and parking slots. Therefore, we computed the expected future state based on the last observed value. For classification purposes, we considered that if a deficit of bicycles or parking slots occurs, the following action should overturn such a deficit. If no deficit is observed, no action is taken. For regression purposes, we considered that if a deficit of parking slots arises, we remove bicycles to reach the minimum number of parked bicycles historically observed at that station after no parking slots were available. Analogously, suppose no bicycles are available at the station. In that case, we consider adding a number of bicycles equal to the maximum number of bicycles observed historically at that station after no bicycles were available.

4.5 Model training

To train and evaluate the models, we performed a ten-fold stratified cross-validation, considering the labels of the classification problem. We used the same folds to train the classification and regression models. For the local models, we considered only the data for the same station we predicted. In contrast, for the global models, we trained the model on the data of the train folds of the station we predicted in addition to all the data available from the remaining stations (see Fig. 4).

For the classification problem, we compared two models: logistic regression and gradient boosted trees[2], training the gradient boosted trees for ten iterations using a multiclass loss function. We also compared two models for the regression problem: linear regression and gradient boosted trees, training the gradient boosted trees for ten iterations and considering the Root Mean Square Error (RMSE) as the loss function.

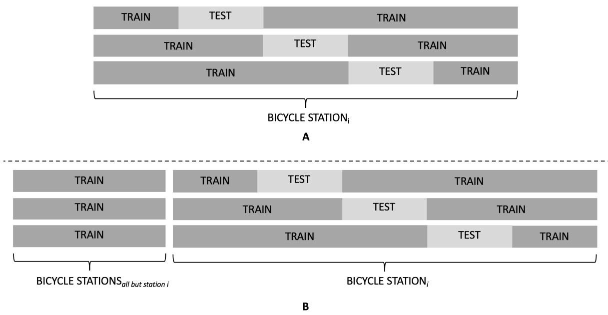

Figure 4: Cross-validation training procedure for (A) local and (B) global models. Given a station *i* for which a forecast is required, local models were trained only on the train folds of that particular station and evaluated on the test folds. Global models, on the other hand, followed the same cross-validation but enriched the train set with data from all the remaining stations.

5 Results and Analysis

We evaluated classification models' discriminative power with the Area Under the Receiver Operating Characteristic Curve (AUC ROC (Bradley (1997))) metric (see Table 2), considering that the metric is invariant to the *a priori* class probabilities, a relevant aspect given the class imbalance observed in the dataset. For the regression problem, we evaluated the models' performance with the Mean Absolute Error (MAE - see Table 3) and RMSE (see Table 4) metrics. While the RMSE metric penalizes large errors, the MAE metric is not sensitive to outliers and therefore provides a better estimate of the models' average performance. The results presented below were computed considering all but the *P+R Barje* bicycle stations.

[2]For the gradient boosted trees, we used the Catboost implementation Prokhorenkova et al. (2018).

#	Feature	Description
1	$mean_5$	The average number of parked bicycles over the last five readings and divided by the station's total capacity.
2	$mean_3$	The average number of parked bicycles over the last three readings and divided by the station's total capacity.
3	$mean_2$	The average number of parked bicycles over the last two readings and divided by the station's total capacity.
4	n-1	The last known number of parked bicycles, divided by the station's total capacity.
5	$mean_5/mean_3$	The average number of parked bicycles over the last five readings and divided by the average number of parked bicycles over the last three readings.
6	$mean_5/mean_2$	The average number of parked bicycles over the last five readings and divided by the average number of parked bicycles over the last two readings.
7	$mean_3/mean_2$	The average number of parked bicycles over the last three readings and divided by the average number of parked bicycles over the last two readings.

Table 1: The features considered to build the classification and regression models.

Model / Forecasting Horizon		20 min.	30 min.	40 min.	50 min.	60 min.
Baseline		0,6658	0,6591	0,6519	0,6443	0,6373
Local	**Logistic Regression**	*0,9901*	*0,9721*	*0,9538*	*0,9369*	*0,9199*
	Catboost Classifier	**0,9939**	**0,9758**	**0,9580**	**0,9418**	**0,9257**
Global	**Logistic Regression**	0,9287	0,9120	0,8959	0,8806	0,8665
	Catboost Classifier	0,9279	0,9114	0,8952	0,8801	0,8654

Table 2: Mean AUC ROC values (higher is better) were obtained across the stations for each forecasting horizon. The mean was computed over the mean AUC ROC obtained from the ten-fold cross-validation performed for each station. Best results are bolded, and second-best are displayed in italics.

Model / Forecasting Horizon		20 min.	30 min.	40 min.	50 min.	60 min.
Baseline		1,3395	1,5189	1,6694	1,8024	1,9247
Local	**Linear Regression**	**0,1607**	**0,6099**	**0,8979**	**1,1177**	**1,3016**
	Catboost Regressor	*0,2444*	*0,6600*	*0,9303*	*1,1370*	*1,3109*
Global	**Linear Regression**	0,8765	1,1722	1,3882	1,5622	1,7125
	Catboost Regressor	0,8642	1,1570	1,3713	1,5396	1,6882

Table 3: Mean MAE values (lower is better) were obtained across the stations for each forecasting horizon. The mean was computed over the mean MAE obtained from the ten-fold cross-validation performed for each station. Best results are bolded, and second-best are displayed in italics.

Model / Forecasting Horizon		20 min.	30 min.	40 min.	50 min.	60 min.
Baseline		2,0567	2,2601	2,4392	2,5998	2,7478
Local	**Linear Regression**	**0,3180**	**1,0519**	**1,4344**	**1,7131**	**1,9393**
	Catboost Regressor	*0,4220*	*1,0941*	*1,4601*	*1,7284*	*1,9460*
Global	**Linear Regression**	1,1319	1,6041	1,9115	2,1503	2,3508
	Catboost Regressor	1,1326	1,5897	1,8952	2,1294	2,3280

Table 4: Mean RMSE values (lower is better) were obtained across the stations for each forecasting horizon. The mean was computed over the mean RMSE obtained from the ten-fold cross-validation performed for each station. Best results are bolded, and second-best are displayed in italics.

We found that all models we created performed better than the baseline. On average, the local models performed best, with the Catboost Classifier leading over the Logistic Regression when predicting the decision-making actions and the Linear Regression leading over the Catboost Regressor when predicting the number of bicycles to be found at the station for a particular forecasting horizon. Nevertheless, when considering global models, we found that the Logistic Regression displayed slightly better performance against the Catboost Classifier in all cases, and the Catboost Regressor outperformed the Linear Regression when measuring MAE and RMSE in all cases, except for RMSE at the time horizon of twenty minutes.

When analyzing the impact of the time horizon on the models' performance, we found that the local models are not strongly affected, displaying an AUC ROC above 0,9 even when predicting an hour ahead. Regression models, on the other side, suffer more pronounced performance degradation. The best model increases the error eight and six times when comparing MAE and RMSE correspondingly, between time horizons of twenty and sixty minutes.

Finally, we were interested in which stations our forecasting models displayed the best and worst performance. We found that local models performed best when predicting decision-making for the *Črnuče* station and the *Zalog* station when forecasting the number of bicycles available. On the other hand, while the worst performance for the regression problem was observed for the *Nama* station, for classification different stations were found depending on the forecasting horizon (e.g., *P+R Dolgi Most* for twenty and thirty minutes, *Lidl Rudnik* for forty and sixty minutes, and *GH Šentpeter Njegoševa cesta* for fifty minutes). When considering global models, decision-making was best predicted for the *Bordarjev trg* station, while the worst performance was observed for the *Studenec* station. For regression, we found the best performance was displayed at the *BS4 Stožice* station. The worst performance for the MAE metric was found at the *Zalog* station. In contrast, for the RMSE metric, the worst performance was split among the *Zalog* (predicting twenty and thirty minutes ahead), *Kolodvorska ulica* (predicting forty minutes ahead), and *Nama* (predicting fifty and sixty minutes ahead) stations. Finally, when looking at the baseline performance for classification, we found that the best performance was observed at the *Živalski vrt* station and the worst one for the *Športni Center Stožice* in all forecasting horizons, except when forecasting sixty minutes ahead, where the worst performance was obtained for the *Mercator Center Šiška*. On the other hand, when considering the regression problem, we found that the best performance was obtained at the *Zalog* station, while the worst one at the *Parkirišče NUK 2-FF* station.

6 Conclusion

Optimizing BSS enables to amplify their use further and minimize the adverse effects on nature. In our work, we trained and compared ML models to predict when rebalancing of the bicycles across stations is required. We evaluated our approach on real-life data from BicikeLJ system in Ljubljana and achieved superior results compared to the baseline methods.

Given that our model can predict bicycle shortage or overflow in advance, this information is helpful for the caretakers of the system. Bicycles could be relocated on time by exploiting other means of transport, thus allowing for higher bicycle use. Another way we can exploit the model's insightful information is to increase the usability of the BSS by incentivizing users to move bicycles to the desired locations, i.e., from stations with no parking capacity to the empty ones.

The main limitation of our study is the real-life implementation, where the relocation of bicycles should consider at least three constraints: (a) must be performed timely, (b) in an environment-friendly way, and (c) be financially efficient. Therefore, our future work will aim to develop a holistic solution while considering the introduction and evaluation of our results in the practical operation of Ljubljana BSS, and the above-mentioned constraints. Furthermore, we plan to test our approach on multiple BSS.

References

Almannaa, M. H., Elhenawy, M., and Rakha, H. A. (2020). Dynamic linear models to predict bike availability in a bike sharing system. *International journal of sustainable transportation*, 14(3):232–242.

Ashqar, H. I., Elhenawy, M., Almannaa, M. H., Ghanem, A., Rakha, H. A., and House, L. (2017). Modeling bike availability in a bike-sharing system using machine learning. In *2017 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS)*, pages 374–378. IEEE.

Ashqar, H. I., Elhenawy, M., Rakha, H. A., Almannaa, M., and House, L. (2021). Network and station-level bike-sharing system prediction: A san francisco bay area case study. *Journal of Intelligent Transportation Systems*, pages 1–11.

Bradley, A. P. (1997). The use of the area under the roc curve in the evaluation of machine learning algorithms. *Pattern Recognition*, 30(7):1145 – 1159.

Breiman, L. (2001). Random forests. *Machine learning*, 45(1):5–32.

Cagliero, L., Cerquitelli, T., Chiusano, S., Garza, P., and Xiao, X. (2017). Predicting critical conditions in bicycle sharing systems. *Computing*, 99(1):39–57.

Masson-Delmotte, V., Zhai, P., Pörtner, H.-O., Roberts, D., Skea, J., Shukla, P. R., Pirani, A., Moufouma-Okia, W., Péan, C., Pidcock, R., et al. (2018). Global warming of 1.5 c. *An IPCC Special Report on the impacts of global warming of*, 1(5).

Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., and Gulin, A. (2018). Catboost: unbiased boosting with categorical features. *Advances in neural information processing systems*, 31.

Rolnick, D., Donti, P. L., Kaack, L. H., Kochanski, K., Lacoste, A., Sankaran, K., Ross, A. S., Milojevic-Dupont, N., Jaques, N., Waldman-Brown, A., Luccioni, A. S., Maharaj, T., Sherwin, E. D., Mukkavilli, S. K., Kording, K. P., Gomes, C. P., Ng, A. Y., Hassabis, D., Platt, J. C., Creutzig, F., Chayes, J., and Bengio, Y. (2022). Tackling Climate Change with Machine Learning. *ACM Computing Surveys (CSUR)*, 55(2):96.

Seo, Y.-H., Kim, D.-K., Kang, S., Byon, Y.-J., and Kho, S.-Y. (2022). Rebalancing docked bicycle sharing system with approximate dynamic programming and reinforcement learning. *Journal of Advanced Transportation*, 2022.

Shaheen, S. A., Guzman, S., and Zhang, H. (2010). Bikesharing in europe, the americas, and asia: past, present, and future. *Transportation research record*, 2143(1):159–167.

Shang, W.-L., Chen, J., Bi, H., Sui, Y., Chen, Y., and Yu, H. (2021). Impacts of covid-19 pandemic on user behaviors and environmental benefits of bike sharing: A big-data analysis. *Applied Energy*, 285:116429.

Sutton, R. S. and Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.

Wu, J. (2020). Challenges and opportunities in algorithmic solutions for re-balancing in bike sharing systems. *Tsinghua Science and Technology*, 25(6):721–733.

Yang, Z., Hu, J., Shu, Y., Cheng, P., Chen, J., and Moscibroda, T. (2016). Mobility modeling and prediction in bike-sharing systems. In *Proceedings of the 14th annual international conference on mobile systems, applications, and services*, pages 165–178.

Yoon, J. W., Pinelli, F., and Calabrese, F. (2012). Cityride: a predictive bike sharing journey advisor. In *2012 IEEE 13th International Conference on Mobile Data Management*, pages 306–311. IEEE.