Universidade de Aveiro
2025

Filipe Gonçalves

Modelos Preditivos para a Redistribuição de Frota num Sistema de Partilha de Veículos

Prediction Models for the Redistribution of the Fleet in a Vehicle Sharing System

DOCUMENTO PROVISÓRIO

Universidade de Aveiro
2025

Filipe Gonçalves

Modelos Preditivos para a Redistribuição de Frota num Sistema de Partilha de Veículos

Prediction Models for the Redistribution of the Fleet in a Vehicle Sharing System

DOCUMENTO PROVISÓRIO

"*Machine intelligence is the last invention that humanity will ever need to make.*"

— Nick Bostrom

Universidade de Aveiro
2025

Filipe Gonçalves

Modelos Preditivos para a Redistribuição de Frota num Sistema de Partilha de Veículos

Prediction Models for the Redistribution of the Fleet in a Vehicle Sharing System

Dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Robótica e Sistemas Inteligentes , realizada sob a orientação científica do Doutor Sónia Gouveia, Professor auxiliar do Departamento de Eletrónica, Telecomunicações e Informática da Universidade de Aveiro, e do Doutor Joaquim Manuel Henriques de Sousa Pinto, Professor auxiliar do Departamento de Eletrónica, Telecomunicações e Informática da Universidade de Aveiro.

Dedico este trabalho à minha família, namorada, amigos e gatos.

o júri / the jury

agradecimentos / acknowledgements

Agradeço aos meus orientadores por todo o apoio durante a realização deste trabalho. Agradeço também à minha família e à minha namorada por me terem aturado durante todo este caminho, aos meus amigos do secundário que estiveram comigo desde antes do início e a todos os meus colegas que me ajudaram durante o meu percurso universitário. Agradeço por fim a todos os meus amigos que não foram mencionados acima e que me ajudaram no que puderam, quando puderam, em especial a Sandra Inês.

Palavras Chave

revisão, análise, Scopus, pesquisa avançada de publicações, partilha de bicicletas, partilha de bicicletas sem estação, primeira/última milha, mudança modal, realocação, problema do rebalanceamento de bicicletas, predição de procura, machine learning, deep learning, dados, infraestrutura, base de dados

Resumo

Os sistemas de partilha de bicicletas são uma alternativa de transporte sustentável, tornando-se mais populares a cada dia que passa e a tendência é de continuar. Desde a sua introdução em 1965, os sistemas de partilha de bicicletas estão categorizados em cinco gerações diferentes, criando inúmeros benefícios como a melhoria da saúde, redução na congestão do trânsito e no combustível consumido. No entanto, problemas como o roubo de bicicletas e o vandalism ou ciclovias inadequadas também ocorrem. Desde estações com demasiadas bicicletas, até à grande procura de bicicletas em outras estações, gerir a frota num sistema de partilha de bicicletas é um dos maiores obstáculos. Este trabalho tem o fim de desenvolver um modelo preditivo para resolver o problema, ao prevêr o número de bicicletas numa certa estação numa certa hora e implementar os seus resultados em duas applicações: uma aplicação móvel para o utilizador puder saber quando terá uma bicicletas livre, e uma applicação de realocação para os redistribuidores saberem quando e onde deverão carregar as bicicletas.

Keywords review, analysis, Scopus, advanced search of publications, bike sharing, dockless bike sharing, first/last mile, modal shift, reallocation, bike rebalancing problem, demand prediction, machine learning, deep learning, data, infrastructure, database

Abstract Bike sharing systems are a sustainable transportation alternative that are becoming more and more popular every day and this trend is expected to continue. Since their introduction in 1965, bike sharing systems are categorized in five different generations, creating many benefits like health improvements, reduced traffic congestion and fuel consumption. However, problems like bicycle theft and vandalism or inadequate bike lanes also occur. From over supplied stations, to higher demand of bikes in different stations, managing the fleet of a bike sharing system is one of its biggest obstacles. This work aims to develop a prediction model to solve the issue, by predicting the number of bikes in a certain station at a certain point in time and implementing the results in two applications: a user mobile application to let the user know whenever a bicycle is free, and a reallocation application to let redistributors know when and where to load the bicycles.

acknowledgement of use of AI tools

Recognition of the use of generative Artificial Intelligence technologies and tools, software and other support tools.

I acknowledge the use of QuillBot Paraphraser (https://quillbot.com/paraphrasing-tool) to help write clearer and more effective sentences.

Contents

List of Figures

List of Tables

List of Code Scripts

Glossary

BSS	Bike Sharing System
DBSS	Dockless Bike Sharing System
GPS	Global Positioning System
BRP	Bike Repositioning Problem
DP	Demand Prediction
RR	Ridge Regression
AR	Adaboost Regression
RFT	Random Forecast Tree
CDA	Community Detection Algorithm
AHC	Agglomerative Hierarchical Clustering
KMC	K-Means Clustering
LDA	Latent Dirichlet Allocation
LSTM	Long Short-Term Memory
ARIMA	Autoregressive Integrated Moving Average
HA	Historical Average
RF	Random Forest Algorithm
LR	Linear Regression
XGBoost	Extreme Gradient Boosting
KNN	K-Nearest Neighbours
SVM	Support Vector Machine
ANN	Artificial Neural Network
RNN	Recurrent Neural Network
ARMA	Autoregressive Moving Average
ARIMAX	Autoregressive Integrated Moving Average with Exogenous Inputs
SARIMA	Seasonal ARIMA
MA	Moving Average
GBRT	Gradient Boost Regression Tree
VAR	Vector Auto-Regression
CART	Classification and Regression Trees
GWR	Geographically Weighted Regression
SVR	Support Vector Regression
KNN	K-Nearest Neighbours
GCNN	Graph Convolutional Neural Network
DTCNN	Dynamic Transition Convolutional Neural Network
DCRNN	Diffusion Convolutional Recurrent Neural Network
LASSO	Least Absolute Shrinkage and Selection Operator
HP	Hierarchical prediction
ER	Error Rate
MAE	Mean Absolute Error
MAER	Mean Absolute Error Rate
MAPE	Mean Absolute Percent Error
SMAPE	Symmetrical Mean Absolute Percent Error
MSE	Mean Squared Error
RMSE	Root Mean Squared Error
RMSER	Root Mean Squared Error Rate
RMSLE	Root Mean Squared Logarithmic Error
PCC	Pearson's Correlation Coefficient
R^2	Coefficient of Determination
CV	Coefficient of Variation

CHAPTER 1

Introduction

1.1 Motivation

Bike sharing refers to the shared use of a bicycle fleet, and has emerged as a sustainable transportation alternative in response to growing concerns about global motorization and climate change. Some of the benefits on the development and use of a bike sharing system Bike Sharing System (BSS) include enhanced mobility, cost savings and lower implementation compared to other modes of transportation, reduced traffic congestion, fuel consumption and greenhouse gas emissions, improved public health and increased environmental awareness [1]. Standard BSSs allow users to rent a bike from one docking station and return it to another, however, not all bike sharing systems require docking stations. Dockless Bike Sharing Systems (DBSSs) provide rapid and flexible mobility, as bikes can be found and parked anywhere. These short-term bikes might feature technologies such as a built-in Global Positioning System (GPS) device that allow operators to track their movements and position, either in a station or during a trip, to prevent theft [2][3]. Since their origin in 1965, in Amsterdam, Netherlands, BSSs have grown exponentially, and most continents now operate them [1]. Their primary travel purposes include work-related activities, leisure across different age groups and residential consumption [4].

However, BSSs have many problems like bicycle theft and vandalism, inadequate bike lines, or the reallocation of both randomly parked bikes in DBSSs and parked bikes from less demanding stations to highly requested stations in BSSs. The literature divides this last obstacle into two categories: the Bike Repositioning Problem (BRP) and the Demand Prediction (DP) problem [5]. The first is solved by creating optimal and efficient truck routes for a vehicle to load bicycles that need repositioning and unload them in high demand stations or areas. The latter is solved by designing, modeling, testing and training a machine or deep learning problem to predict the number of bikes in a station or area at a certain time, using the trip data from a BSS, and compare it to other artificial intelligence models using different evaluation metrics.

The work done in this dissertation is to solve the latter problem with the goal of implementing the results into two different applications: a mobile application to allow a user to easily find a bike to start a trip, and a reallocation application to alarm the reallocation team on where bikes need to be reallocated.

1.2 Objectives

As mentioned in the previous section, one problem of BSSs is the reallocation of parked bikes from one station to another, new bikes into different stations, or randomly parked bikes to different areas. From this, it is possible to define the main objectives of this dissertation, outlined as follows:

- **Compare different machine and deep learning models**: By analyzing different time-series models, it is possible to determine which is the most suitable machine or deep learning model to predict the number of available bikes in a station at a certain point in time;
- **Build, train and test the model**: By developing the chosen machine or deep learning model, and tuning the hyperparameter, it is possible to create the most efficient and accurate model, while comparing the results to other state of the art models;
- **Implement the model in diverse applications**: By dynamically introducing the results of the predictions in a reallocation application, the truck drivers will know when a high demand station will be under supplied, and vice-versa, while implementing the results in a user mobile application, a user will have greater insight into when a bike will be available in a specific station.

1.3 Outline

This section offers a brief summary of each chapter and the arrangement of the dissertation's content.

In addition to this introductory chapter, the document is divided into five more chapters:

- **Chapter 2 - Search and Screening Methodology**: The methodology of the search for relevant papers is presented, focusing on the inclusion of papers based on a theme system, and the exclusion of papers based on exclusion criteria;
- **Chapter 3 - Literature Review**: A review on the current state of art of BSS is analyzed, focusing on their problems, potential solutions and their implications, with a variety of examples of their successes and failures, as well as an analysis on the theoretical background of demand prediction's machine and deep learning models;
- **Chapter 4 - Preliminary Work**: A summary of the work realized of the dissertation is described, including the state of the art, theoretical background and an infrastructure analysis.

- **Chapter 5 - Work Plan**: The plan for the next segment of the dissertation is presented, stating the tasks needed and an estimation on when they will be started and completed.
- **Chapter 6 - Conclusion**: The conclusions drawn from the literature review and preliminary work are discussed, as well as providing insights into the future work.

CHAPTER 2

Search and Screening Methodology

This chapter presents the methodology used in the search for relevant scientific publications. Briefly, the comprehensive search was conducted in the Scopus database, followed by a systematic screening process and an in-depth analysis of the selected literature.

2.1 Complete Search

The first step in any selection process is to identify the best criteria. Any variation of these arguments can lead to disparate outcomes, from the many methods of identifying a BSS to the type of document to screen or the time frame for the review. The terms "bikesharing", "bike sharing", and "bike-sharing" were defined as key expressions to appear in the title, keywords or abstract. Additionally, the papers must be published and written in English, and they must be published before 2025. The full script for the complete search is as follows: `TITLE-ABS-KEY ("bike-sharing" OR "bikesharing" OR "bike sharing") AND LANGUAGE (english) AND PUBYEAR < 2025`.

This search produced more than 2,700 documents with distribution presented in Figure 2.1, of which 60% were articles and 30% were conference papers.

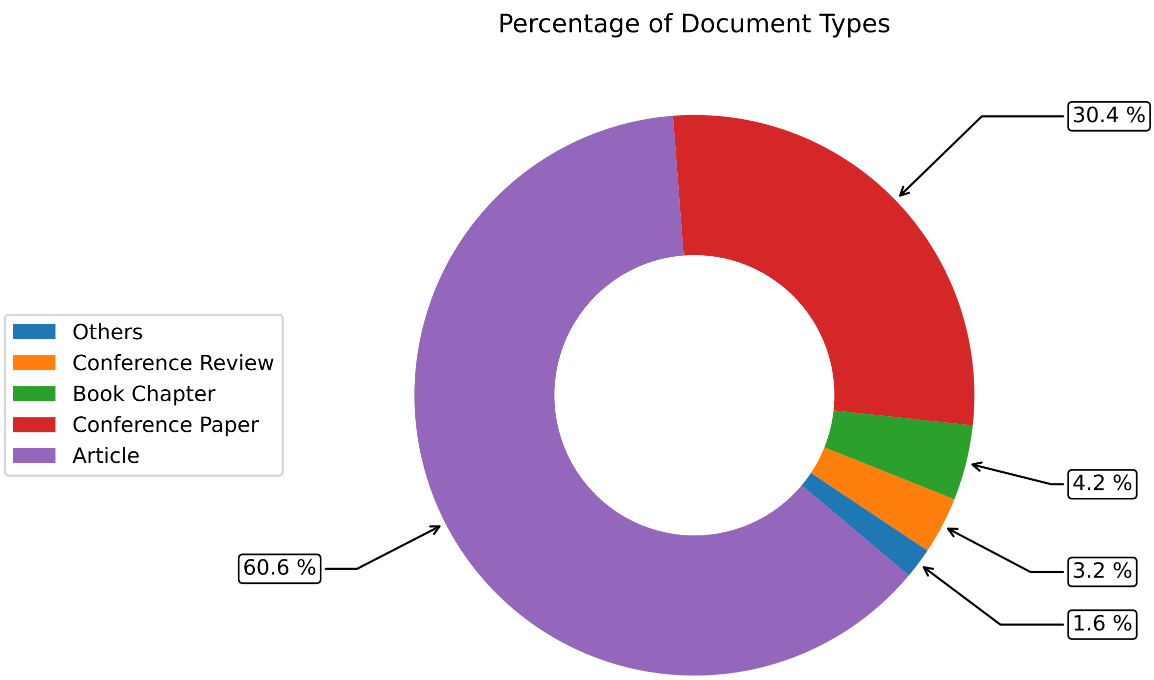

Figure 2.1: Types of the documents

2.2 Advanced Search

An advanced search was carried out, using the same key expressions and time frame, but restricting the source of documents to only journals, with the query: `TITLE-ABS-KEY ("bike-sharing" OR "bikesharing" OR "bike sharing") AND SRCTYPE (j) AND LANGUAGE (english) AND PUBYEAR` $<$ `2025`.

This search resulted in nearly 1,800 documents, that were then manually screened based on the title and abstract and divided into two different relational tables, one for the included papers aggregated using a theme system, which yielded a total of 1,316 papers, and another for the excluded papers based on a exclusion condition system.

Figure 2.2 depicts the number of total and included documents published in each year.

With the number of published papers rising at an almost exponential rate each year, BSSs are expanding like never before. It is evident that there is much to learn about this topic: from a handful of papers prior to 2011 to nearly achieving 300 papers in 2021, with the number of publications remaining above 250 every year after, and even though BSSs were widely used before the start of the decade, more than 90% of all documents were published after 2016.

Across all 1,316 papers, the inclusion selection was implemented using a theme system, in which one paper is outlined by the most relevant theme for easier filtering, listed as follows:

- Case Study: Any general study on bike sharing - the economy, the impacts, the problems, how to promote a BSS, comparative analysis of different types of BSSs, studies on BSSs, and so on;
- Rebalancing: Any study on the bike rebalancing problem, demand prediction mod-

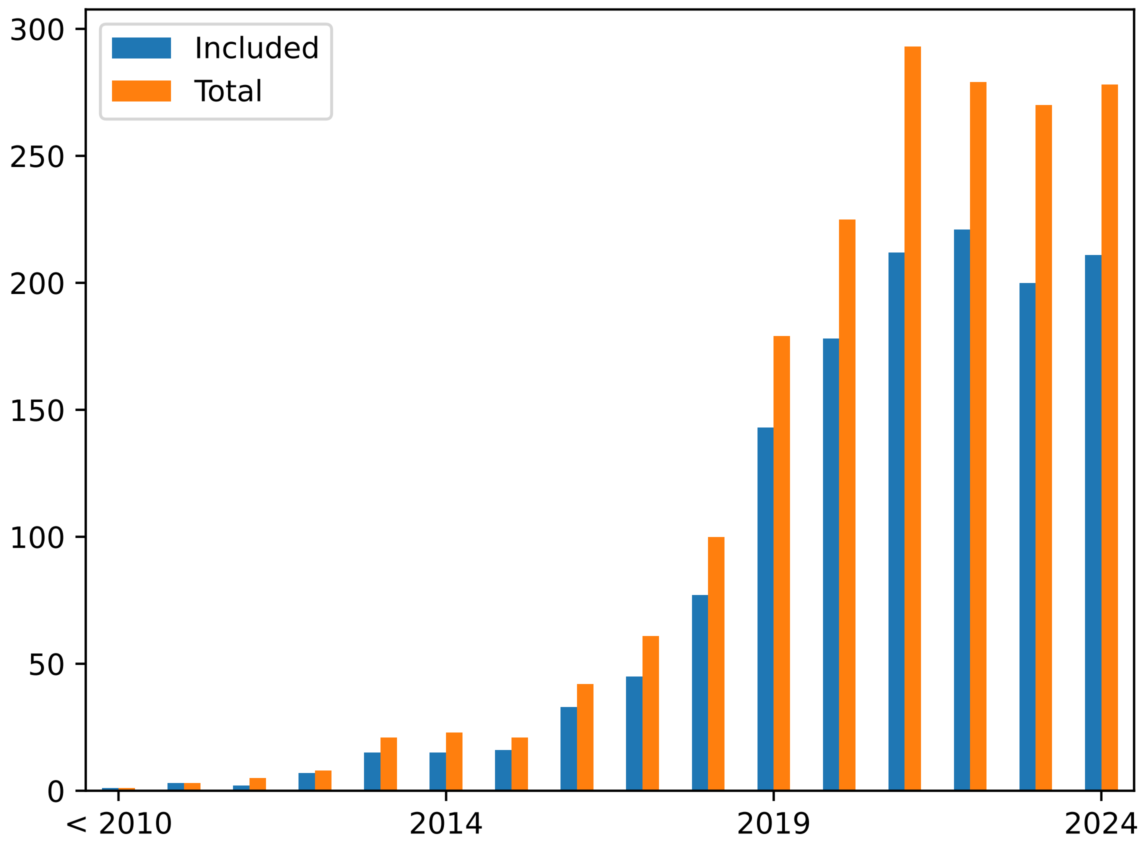

Figure 2.2: Number of Total and Included Papers per Year

els, algorithms for the location of the stations, travel pattern recognition for route recommendations and spatio-temporal models for BSSs;

- Examples: Any study on a specific city (studies on countries could be classified based on the impacts, problems, etc.), which could be of the failure or success of a given BSS;
- First/Last Mile: Any study on the correlation between a BSS and any transportation method (metro, bus, taxi, etc.), where the user's journey begins or ends on a BSS;
- Sustainability: Any research on the green impact of BSSs.

Similarly, the exclusion criteria for the excluded papers is listed below:

- Mobility: Any study on mobility as a service like taxi demand or vehicle travel time, and mobility sharing like car or scooter sharing;
- Not Focused on Bike Sharing: Any study that uses bike sharing as an example or as a dataset to a problem outside the scope of BSSs;
- Error: Any erratum study.

Figure 2.3 shows the percentage of included and excluded papers based on the different themes, in green, and exclusion conditions, in red, providing an analysis of how in-depth the papers are becoming. The majority of the included papers analyze either different case studies or the rebalancing problem, and consequently the discussion and evolution in these areas is far more developed compared to studies on sustainability, which represent 4% of the total number of documents. Moreover, almost 23% of all papers were excluded from the screening process as most of them are studies on mobility as a service, shared mobility or studies that reference BSSs as an example or use data from BSSs to solve their problems. Nonetheless, more than three quarters of all the documents were selected for the next process, meaning that the majority of the literature can be included in any type of analysis of BSSs.

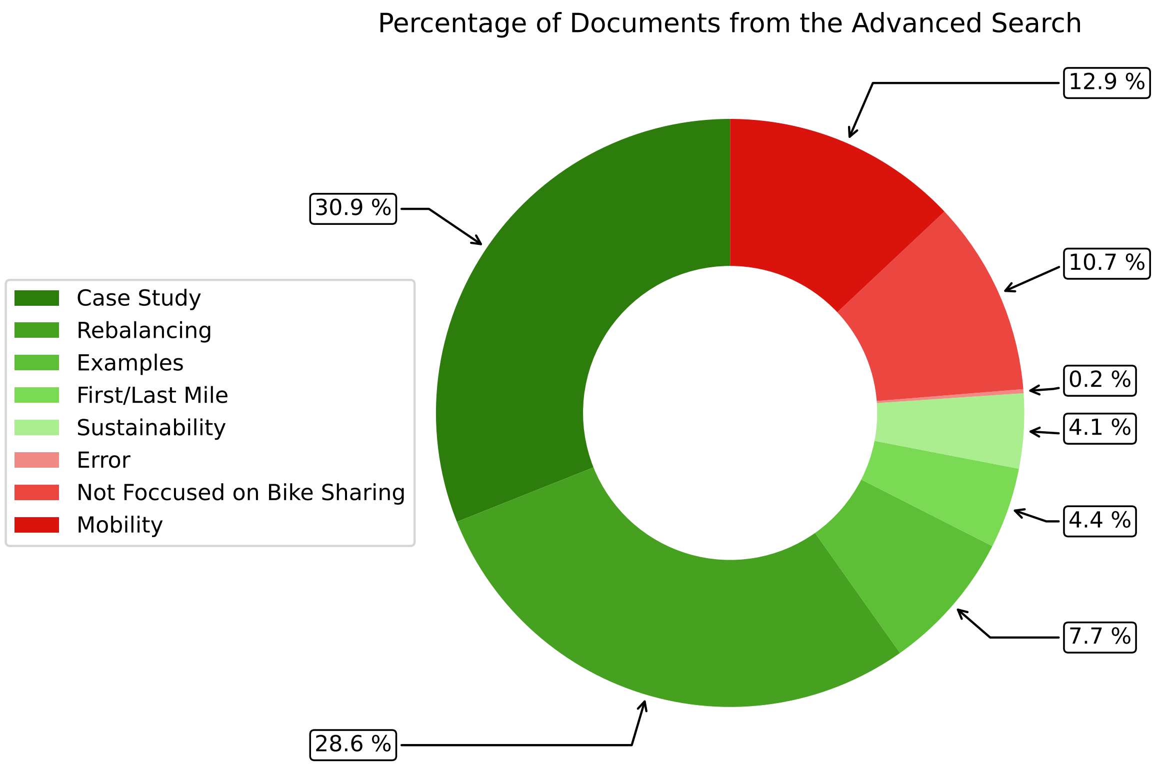

Figure 2.3: Themes and Conditions for the 1316 Documents

2.3 Paper Screening

1,316 papers resulted from the advanced search. While there is a variety of useful information about BSSs, it is hard to do an in-depth analysis of each one. However, filtering papers by number of citations, publication year, or even both has certain problems: Filtering by number of citations means that recent papers that have yet to receive as many citations will be ignored; Filtering by year implies that the number of citations is meaningless, which is false; Filtering by year and number of citations means that if we only get the best papers each year, papers from later years will be excluded. The solution is to use both the publication year and the number of citations to calculate the *impact* (2.1) of the papers.

$$impact = \frac{number\_of\_citations}{number\_of\_years\_since\_publication} \tag{2.1}$$

The *impact* will look into the mean number of citations per year published, which solves the limitations of filtering based on only the number of citations or the number of published years and ensures that we get the best papers possible.

This process reduced the number of included papers to approximately 500, with almost 340 of these papers discussing and analyzing the rebalancing problem.

CHAPTER 3

Literature Review

This chapter addresses the evolution of bike-sharing programs worldwide, their problems, potential solutions and their implications, as well as a variety of examples of their successes and failures. This will be followed by an overview on the reallocation problem, with an analysis of the current literature and theoretical background on demand prediction's machine and deep learning models.

3.1 State of the Art

Previous studies summarized and classified bike sharing systems into four generations [1]:

- **First Generation**: White bikes (free bike systems) introduced in Amsterdam in the 1965;
- **Second Generation**: Coin-deposit systems, starting with Bycyklen in Copenhagen in the 1995;
- **Third Generation**: IT-based systems, the first of which appeared in Rennes, France, in the 1998;
- **Fourth Generation**: Focusing on demand-responsive, multi-modal systems, in the 2009.

The first generation relied significantly on the use of free bicycles that were often randomly parked throughout the city. Without any knowledge on where exactly the bicycles where, on whether they were intact or damaged, on whether users sold them, took them home or vandalized them, many bike sharing systems were left ignorant on their fleet. This resulted in the failure of Cambridge's Green Bike Scheme, with nearly 300 stolen bicycles [6] and the failure of the Amsterdam's White Bikes, which had some of their fleet thrown into channels [7].

The second generation is best known for its docking system, which allowed bikes to be parked and locked in stations. Bicycles were then unlocked with a coin deposit of roughly 3$, which was refunded upon return. The first of this generation is the Bycyklen in Copenhagen, which still continued to operate in 2012 with over 2,000 bicycles and 110 stations [1]. In this system, users could have their bicycles for indefinite time, as their only loss was the coin deposited in the docking station, so they frequently took bicycles for long periods of time, continuing the cases of theft and vandalism [8].

The third generation gained popularity by incorporating innovative technologies for bicycle tracking, parking and pickup. The IT-based systems introduced bicycles with locks, paid memberships and an anti theft program that required users to show their ID to ride the vehicles. In 2007, Paris launched its biggest bike sharing system to date, named Vélib with over 20,000 bicycles [3] and 20 million trips between 2007 and 2008, making it the largest in Europe [1].

The fourth generation would eventually include electric bicycles, more efficient and flexible stations, a simpler user interface, and some form of a redistribution system. The Montreal's BIXI program marks the start of this generation, with over 5,000 bicycles and 400 stations. It features solar-powered mobile stations that can be relocated based on usage patterns and pick-up or drop-off agglomeration, as well as GPS integration into the bikes, which reduces theft and vandalism compared to previous systems and allows for easier bike redistribution [1].

While the concept, design and technologies of global bike sharing systems have been evolving since well before the century started, it is not until 2007 that we can visually see the results of these improvements, shown in Figure 3.1 until the second quarter of 2012.

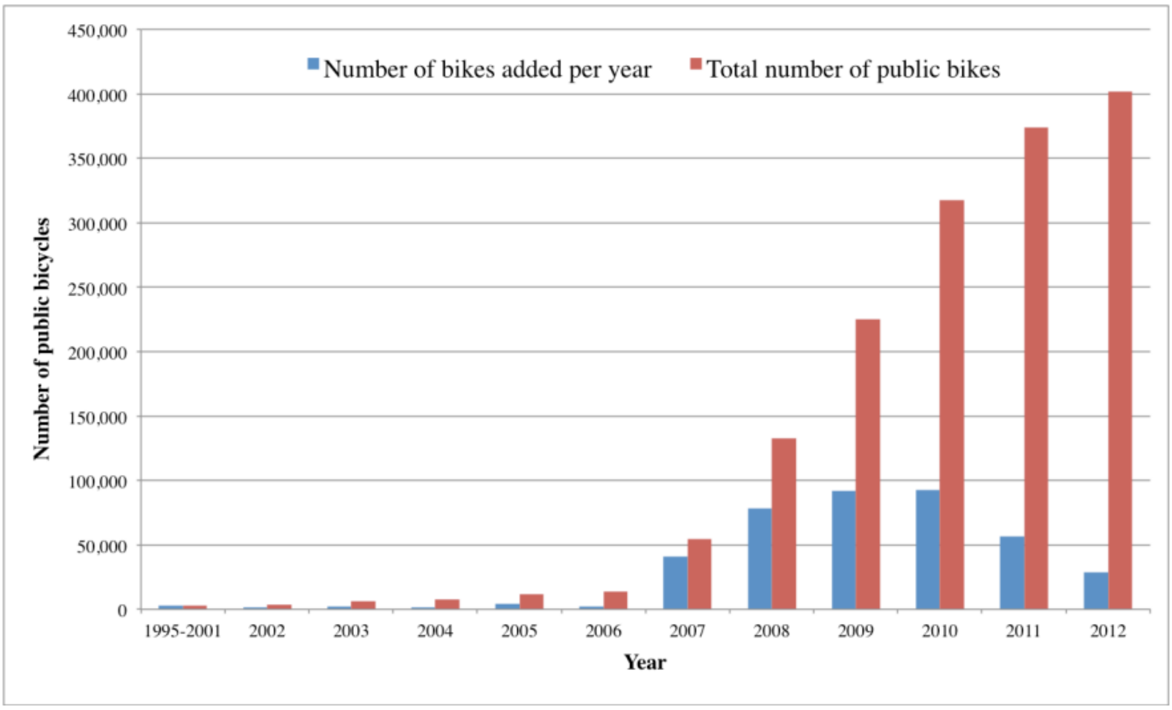

Figure 3.1: Global growth in bike share programs. [3]

In 2005, China entered the world of bike sharing systems with the Beijing program. From then and until 2012 China had twelve cities with formal public bike sharing systems, totalling over 180,500 bikes [9]. In 2011, Wuhan and Hangzhou had world's largest systems with 70,000 and 65,000 bicycles, respectively [10].

The sudden implementation of many bike sharing systems is rooted in the decline in bicycle utilisation since the 1990s [11]. The economic growth, longer trip distances and deteriorating cycling environments were the major factors behind said decline, as in the 1970s China was called "Kingdom of Bicycles" for its reliance on bicycle usage, and later on averaging 113 bikes per 100 households [11].

There are many more examples of bike sharing systems being studied around the world, whether it is to just analyse its implications [12][13], impacts [14], desires [15], demand [16] and even ridership influencing factors [17], analyse user's behaviour and adoption [11][18], create machine learning models [19], compare different bike sharing approaches, or to just analyse and study a new case study [20][21][22][23] [24].

Nonetheless, the increased demand and thirst for evolution and improvements in bike sharing systems led to the development of new technologies and different approaches to the problem. The fourth generation of bike sharing systems created more efficient stations for dropping off and picking up bicycles, but it also introduced a new technology [8]: the dockless station. This innovation makes it easier for any user to complete or begin a trip, since the bicycles can be found anywhere and the users can access them instantly without the responsibility of returning them to a specific dock [25]. A reallocation team is then required to organize all the bikes into different parking areas. This technological advancement marked the start of the fifth generation, the dockless bike sharing systems [26].

The implementation of dockless bike sharing systems, also known as free-floating bike sharing systems in China, increased cycling in urban areas, with significant reduction in car usage and carbon emissions [27]. In 2016, both Mobike and Ofo's dockless bike sharing systems were introduced in China's first tier cities, Beijing, Shanghai, Guangzhou and Shenzhen, and in just one year their popularity skyrocketed. The previous four cities quickly grew to 200 cities, with over 23 million bikes and 221 million users [28].

Additionally, depending on the transport mode substituted by the use of bike sharing systems, defined as modal shift [8], dockless or standard programs could be of great benefit. Kun Gao et al. [29] indicates that depending on the mode of transportation substituted, the modal shift can be beneficial or not. Replacing a private car, resulting in fewer car trips, reduces energy consumption and exhaust emissions during the trip; Replacing transit trips will not be of benefit, as schedules won't change; Replacing walking may even impose negative impacts on dockless bike sharing systems.

Nonetheless, utilizing these systems to commute between two modes of public transport is very common: In Beijing a significant percentage of shared bikes were found active near bus stops and metro stations [25]; A study on how car congestion near subway stations can affect bike utilization [30] concluded in favour of the transportation mode substitution; From a survey made on car drivers [31], travel distance didn't matter, and users were willing to substitute their mode of transport.

However, when compared with standard systems, dockless bike sharing programs have significantly more theft, misuse and vandalism-related problems [32]. Bikes can be parked anywhere, occupying urban space, impacting traffic [28] and limiting access to facilities, pedestrian and cycling lanes and bus stops [29]. The primary challenge, however, is a direct consequence of these problems: The reallocation of the bicycles, which emit more greenhouse gases the more disperse are the bikes [28]. In standard bike sharing systems, the rebalancing problem is also a big issue as some stations can be under supplied and have high demand, leading to the failure of China's standard and Singapore's dockless bike sharing systems in 2019 [33].

3.2 Theoretical Background

Whether it is simply inserting new bikes into stations, reallocating existing bikes into new stations, moving bikes from different stations or, in dockless bike sharing systems, repositioning the randomly parked bikes into specific areas, this process requires not only manpower, but also some kind of system to manage which bikes are repositioned to which stations, if any. From the last screening process it was possible to divide this problem into two different categories:

- **BRP**: Calculation of the optimal truck routes and the loading and unloading of bikes at each station [5];
- **DP**: The prediction of the number of bikes left in a station at a certain point in time.

There are multiple studies that try to create efficient truck routes to minimize the greenhouse gas emissions [5] or user dissatisfaction [34] while minimizing traveling time and/or distance between stations. Others try to incentivize the users to reallocate the bikes themselves, with lower expected cost to end a trip in a station with less bikes or higher demand, comparatively to another station less utilized or with more locked bikes [35][36]. The problem also lies in the number of bikes loaded and unloaded at each station [5], which is dependent on the route the truck will take.

Many of these studies focus on achieving a solution based on mathematical algorithms, with ordinary differential equations [37] and/or decision processes [38][39][40], or mathematical optimizations [41][42][43][44][45]; Applying some kind of heuristic [46][47][48]; Basing the solution around the stations, using their location [49][5] and/or grouping them into clusters [46][45]; and many more.

In comparison, the demand prediction problem is more straightforward but equally complex. Fixing it is crucial for effective bike repositioning, as stations can't be left empty while others are full [50].

Lei Lin et al. [50] divided the prediction models in the literature in three groups:

- **City-level**: Focused on predicting the total bike usage for an entire city;
- **Cluster-level**: Focused on grouping stations into clusters;
- **Station-level**: Focused on predicting demand for each station.

The first group, while it simplifies the issue, doesn't contribute in solving the rebalancing problem. Furthermore, it doesn't utilize most of the data extracted from a bike sharing system, like trip duration, origin and destination [50]. An example is the study in [51], which made predictions for the next day using Capital Bike Sharing system's data, using algorithms like Ridge Regression (RR), Adaboost Regression (AR), Support Vector Regression (SVR), Random Forecast Tree (RFT) and Gradient Boost Regression Tree (GBRT).

The second group creates clusters of stations based on geographical locations and/or temporal demand patterns, assuming stations can be correlated and that the total demand of these stations can be predicted as a cluster [50]: Stations near a residential area might behave similarly between them, as well as stations near shopping centres or stations near public transport stops. Some examples of cluster-level predictions are: Community Detection Algorithm (CDA) and Agglomerative Hierarchical Clustering (AHC) [52], and K-Means Clustering (KMC) and Latent Dirichlet Allocation (LDA) [53].

The third group is much more challenging to implement [50]. Some studies implement linear regression models to predict monthly [54] or hourly [55] rentals, while others go further with neural networks [50][56] or and other machine and deep learning models.

However, these groups were devised with only station-based systems in mind. Travel patterns and trip frequency provide valuable insights in understanding the demand of these systems, but nothing can be said about the demand on dockless bike sharing systems [57]. Without a station to forecast the number of bikes at a certain time, it becomes difficult to predict the demand in a singular place.

The demand prediction is mostly done by modelling the problem using a machine or deep learning algorithm. Table 3.1 depicts the ten most used artificial intelligence models, either to solve the problem or as baseline to compare to original models and the most used evaluation metrics to score the models, for the 50 most impactful papers in demand prediction.

Since the solution is to predict the number of bikes at a certain point in time, a time-series forecasting model is necessary. When modeling this type of data in bike sharing systems, the ANN [57] is one of the most common algorithms used, however it doesn't account for temporal dependencies in the model structure. By recurrently connecting hidden layers at different timestamps, RNN [57][58][59][56] could overcome the limitation, but it is still not fit for the time-series data with long time lags [57]. To solve the problem, the LSTM is used in several studies, with neural networks to fit the data of dockless bikes [57], with an irregular convolutional layer to predict bike demand among similar urban areas [60], with a convolutional neural network for feature extraction on input data combined with the support sequence prediction of LSTM [61], and with spatial and temporal usage patterns to predict the short-run demand for station based systems [62]. It was also used as baseline comparator in many other studies [50][56][63][64][59][58].

Just like ANN, the ARIMA is commonly used for forecasting time-series, applied widely in traffic prediction [65]. Its performance can be significantly influenced by model tuning, and several variants can also generate better results [63], like Autoregressive Moving Average (ARMA) which is commonly used for understanding and predicting future values in a time series [66][67][68], and the Seasonal ARIMA (SARIMA) and Autoregressive Integrated Moving Average with Exogenous Inputs (ARIMAX) [69]. It is also used as baseline comparator in many other studies [60][58].

In regards to HA, it uses the historical average demand for prediction. Predicting the

Table 3.1: Demand Prediction Models by Number of Papers

Model	# Papers	Evaluation Metrics
Long Short-Term Memory (LSTM)	11	RMSE, MAE, MAPE, MSE, PCC, SMAPE
Autoregressive Integrated Moving Average (ARIMA)	11	RMSE, MAE, MAPE, R^2, RMSER, MAER, RMSLE, ER
Historical Average (HA)	9	RMSE, MAE, MAPE, R^2, PCC, RMSLE, ER
Random Forest Algorithm (RF)	9	RMSE, MAE, R^2, CV, MAE, RMSE, RMSLE
Linear Regression (LR)	9	RMSE, MAE, R^2, CV, RMSER, MAER
Extreme Gradient Boosting (XGBoost)	6	RMSE, MAE, MAPE, PCC
K-Nearest Neighbours (KNN)	5	RMSE, MAE, MAPE, R^2
Support Vector Machine (SVM)	5	RMSE, MAE, MAPE, R^2, CV
Artificial Neural Network (ANN)	4	RMSE, MAE, MAPE, MAE, RMSE, RMSLE
Recurrent Neural Network (RNN)	4	RMSE, MAE, MAPE, R^2, PCC, SMAPE, MSE

travel demand for a specific day at a specific time will result in the average from all past values in that specific day and time in the training dataset[63]. This model usually only appears as a comparator to other models[56][64][57], as both HA and its variant Moving Average (MA) are naive but competitive models[70].

Furthermore, RF can be used to overcome the problem of a model containing large amounts of predictors, becoming more complex and ovefitting the data. The RF creates an ensemble of decision trees and randomly selects a subset of features to grow each tree [71]. It was applied as baseline comparator [72] or as the solution model using seasonal changing data [73].

Another model used is the SVM [57], that was developed to seek the finest hyper-plane classification of data, which is transformed into a high-dimensional feature space where it can be linearly segregated to differentiate two groups that are not linearly separable [74]. It's used as baseline comparator [74], but an important branch has been applied widely in traffic flow prediction, SVR [64], or as a baseline comparator [58][73].

In comparison, LR is the simplest method, taking one or more X features in the input and outputting the number of bikes predicted [74]. Many regressions have been used as comparator models like the GBRT [66], Vector Auto-Regression (VAR) [56], Classification and Regression Trees (CART) [73], Geographically Weighted Regression (GWR) [75], as well as the standard LR [69].

Others like XGBoost [57][50][63][56] that are based on an implementation of gradient boosted decision trees, and KNN [66][75][73] which predicts based on it-s K closes neighbors, are mostly used as baseline comparators.

More models have been considered, including, convolutional and recurrent networks (Graph Convolutional Neural Network (GCNN) [50], Dynamic Transition Convolutional Neural Network (DTCNN) and Diffusion Convolutional Recurrent Neural Network (DCRNN) [56]); Least Absolute Shrinkage and Selection Operator (LASSO) [50] and Hierarchical prediction (HP) [66]. Most of these models, which are implemented just a handful of times, are largely used to compare to other models.

These studies use various metrics to evaluate model performance, including error metrics such as Error Rate (ER), Mean Absolute Error (MAE) and its rate (Mean Absolute Error Rate (MAER)), Mean Absolute Percent Error (MAPE) and Symmetrical Mean Absolute Percent Error (SMAPE). Other error metrics include Mean Squared Error (MSE), Root Mean Squared Error (RMSE) and its rate (Root Mean Squared Error Rate (RMSER)) as well as Root Mean Squared Logarithmic Error (RMSLE). Association metrics, such as Pearson's Correlation Coefficient (PCC) and Coefficient of Determination (R^2), as well as variation metrics (Coefficient of Variation (CV)) are also used to assess model performance.

MSE and RMSE are mostly applied when large errors must be penalized, as they are highly sensitive to outliers, however RMSE is widely used in measuring the error of regression models [56]. They are calculated using the following equations, 3.1 and 3.2:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(x_i - y_i)^2 \tag{3.1}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - y_i)^2} \tag{3.2}$$

MAE and MAPE are used when all the errors are treated equally. They are calculated using the following equations, 3.3 and 3.4:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|x_i - y_i| \tag{3.3}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}|\frac{x_i - y_i}{x_i}| \times 100 \tag{3.4}$$

MAE is similar to ER, however MAE is used for regression tasks, while ER is mostly used for classifications tasks where correctness matters. The R^2 is useful for model comparison and variance explanation, it helps understand how good the model fitted the data, as a R^2 value of 1 indicates a perfect model.

The MAE is the average absolute error, which can better reflect the actual situation of the predicted value error. The RMSE is more sensitive to outliers owing to the degree of change in the evaluation data. [76]

Only looking at the evaluation metrics, however, when comparing two different models is not recommended. Some studies predict the number of bikes after 10, 15, 20 or 30 minutes [57] while others only predict after one hour [55], or one month [54], or even in a specific time period [50]. The features used while modeling the solution will most likely be different from one model to another, some might use weather data in their predictions [77], while others use the stations location [50]. The data used to train and test the model is usually different from study to study, being impossible to conclude from two different datasets which model is better.

Choosing from the many existing models is one of the biggest problems in the demand prediction problem. Each use case has different problems at hand, using different datasets and features, and as such creating a model, training and testing is completely subjective.

CHAPTER 4

Preliminary Work

This chapter presents the preliminary work carried out during the first half of the dissertation paper, by detailing each component.

4.1 Systematic Search Scripting

4.1.1 Advanced Search

The search methodology presented in Chapter 2 was executed following the listed plan:

- From the Scopus database for scientific publications all documents' citation information - author(s), document title, year, number of citations, source and document type, DOI and abstract - was downloaded into a text file;
- The citation information in the downloaded file was translated to a table of documents and saved into a *.csv* file, to maintain order and to allow for easier selection and filtering;
- Read the saved data from the *.csv* file and manually selected papers to either include in the screening process by organizing them into different themes, or to exclude by organizing them into different exclusion conditions.

The translation of the information was executed using a script 1 to read all the lines in the file and organize them into different rows on the *.csv* file.

Lastly, the manual selection was conducted using a script 2 to iterate through all the documents, enumerating the title, number of citations, authors, link and abstract, and waiting for the user to either input an "enter" to include the paper and select the theme, or input any other letter to exclude the paper and select the exclusion condition. The script also saved the included and excluded papers in their specific *.csv* file and only iterated over documents that weren't in these files, keeping an offset and letting the user end the program whenever he desired while always saving the last changes made.

```python
import pandas as pd

try:
    articles = pd.read_csv("articles.csv").drop(columns=["Unnamed: 0"])
except:
    file_articles = []
    with open("scopus.txt", "r+") as f:
        f.readline()    # remove source
        f.readline()    # remove timestamp
        f.readline()    # remove newline

        a = {}
        countline = 0
        title = False

        for x in f:
            if (x == "\n"):
                if countline >= 1:
                    countline = 0
                    file_articles.append(a)
                    a = {}
                else:
                    countline += 1
            else:
                if x.startswith("AUTHOR FULL NAMES:"):
                    a["author"] = x.split(": ")[1].replace("\n", "")
                if ", Cited" in x:
                    a["year"] = int(x[1:5])
                    a["number_citations"] = int(x.split(" ")[2])
                if x.startswith("http"):
                    a["link"] = x.replace("\n", "")
                if title:
                    a["title"] = x.replace("\n", "")
                    title = False
                if x[0].isdigit():
                    title = True
                if x.startswith("ABSTRACT: ") == 1:
                    a["abstract"] = x.replace("\n", "").replace("ABSTRACT: ", "")
                if x.startswith("DOCUMENT TYPE") == 1:
                    a["type"] = x.replace("\n", "").replace("DOCUMENT TYPE: ", "")
                if x.startswith("SOURCE") == 1:
                    pass

    articles = pd.DataFrame.from_dict(file_articles)
    articles.to_csv("articles.csv")
```

Code 1: Translation of the text file into *.csv* data

```python
try:
    for index, row in articles.iterrows():
        # already included
        if not artciles_reviwed.empty and \
                artciles_reviwed[artciles_reviwed["title"] == row["title"]]["title"].count() != 0:

            data = artciles_reviwed[artciles_reviwed["title"] == row["title"]].iloc[0].to_json()
            articles_filtered.append(json.loads(data))
            continue

        # already excluded
        if not artciles_discarded.empty and \
            artciles_discarded[artciles_discarded["title"] == row["title"]]["title"].count() != 0:

            data = artciles_discarded[artciles_discarded["title"] == row["title"]].iloc[0].to_json()
            articles_removed.append(json.loads(data))
            continue

        print("-------------------------------------------------------")
        print(f"Index: {index+1}/{total_number} - {math.floor((index+1)*100/total_number)}%")
        print(f"\tAuthors: {row['author']}")
        print(f"\tYear: {row['year']} --- Impact: {round(row['number_citations']/(2024-row['year']+1), 0)}")
        print(f"\tTitle: {row['title']}")
        print(f"\tAbstract: {row['abstract']}")

        x = input()
        r = row.to_dict()

        # only enter
        if x == "":
            print('\n'.join(["\t" + str(i) + " - " + list(themes_dict.keys())[i] \
                             for i in range(len(list(themes_dict.keys())))]))
            temas = input("\tTemas: ")

            if temas.isdigit() and int(temas) < len(list(themes_dict.keys())):
                th_sp = [list(themes_dict.keys())[int(temas)]]
            else: th_sp = temas.split(", ")

            for i in range(len(th_sp)):
                if "Examples" in th_sp[i] and i+1 < len(th_sp):
                    th_sp[i] += " " + th_sp[i+1]
                    th_sp.pop(i+1)
                    break
            r["theme"] = th_sp
            articles_filtered.append(r)
        else:
            print('\n'.join(["\t" + str(i) + " - " + list(conditions.keys())[i] \
                             for i in range(len(list(conditions.keys())))]))
            cond = input("\tCondição de Exclusão: ")

            if cond.isdigit() and int(cond) < len(list(conditions.keys())):
                r["conditions"] = list(conditions.keys())[int(cond)]
            else:
                r["conditions"] = conditions
            articles_removed.append(r)
        print("-------------------------------------------------------")
        print()

except KeyboardInterrupt:
    print("\n\n")
    pass
```

Code 2: Manual Selection of Papers

4.1.2 Screening Process

The screening process reduced the number of included documents to almost 500 papers based on the *impact* of each paper. However, the *impact* of the papers is only fair when the papers belong to the same theme. Consequently, the screening process looked not only at the *impact* of the papers but also at which theme the document is inserted. Papers on "Rebalancing" are more relevant to the objectives of this dissertation, and as such, most of the papers included belong to this theme. On the other hand, it is not relevant to include all the papers on "Examples" or "Sustainability". For this purpose a similar script as the one used in the manual selection of papers (2) was used. The program would only need an "enter" to include the paper and any other letter to exclude the paper, saving all the included documents in a new *.csv* file.

4.2 Literature Review Scripting

The screened documents were sorted according to the *impact* of each paper in order to review the state of art and conduct a theoretical background described in Chapter 3. For easier access to all the documents a script was created to search for documents based on their theme or their name 3.

```python
import pandas as pd

articles_selected = pd.read_csv("screening.csv").drop(columns=["Unnamed: 0"])

while True:
    print("1- Pesqusiar por Tema\n2- Pesqusiar por Título\n3- Pesqusiar Todos\n4- Terminar")

    x = input("Opção: ")
    count = 1

    if int(x) == 1:
        tema = input("Tema: ")
        articles = articles_selected[articles_selected["theme"].str.lower().str.contains(tema.lower())] \
                   .sort_values(by="impact", ascending=False)

    if int(x) == 2:
        title = input("Título: ")
        articles = articles_selected[articles_selected["title"].str.lower().str.contains(title.lower())] \
                .sort_values(by="impact", ascending=False)

    if int(x) == 3:
        articles = articles_selected.sort_values(by="impact", ascending=False)

    if int(x) == 4:
        break

    for i, r in articles.iterrows():
        print("-------------------------------------------------------")
        print(f"Index: {index+1}/{total_number} - {math.floor((index+1)*100/total_number)}%")
        print(f"\tAuthors: {row['author']}")
        print(f"\tYear: {row['year']} --- Impact: {round(row['number_citations']/(2024-row['year']+1), 0)}")
        print(f"\tTitle: {row['title']}")
        print(f"\tAbstract: {row['abstract']}")
        input("Next (Enter)")
        count += 1
```

Code 3: Automatic Search of Papers

4.3 Infrastructure

BSSs incorporate more than just bikes, stations, truck routes and demand prediction models. The development of these programs demand an enormous amount of effort, dedication and work.

First, many dashboards must be implemented, filled with relevant information like the number of bicycles in use, parked and in maintenance, the number of stations in the system and their occupancy, and the number of users in the system and their trip record. All this information is required for the operator's backstage work, since they need to know all the problems the system has and tools to solve them. As some bikes come with an in-built tracking device, it is critical to have a map that indicates the position and state of the bicycles, as well as metrics and statistics to understand if the system is working accordingly or not.

Additionally, with ever growing technology and new "intelligent" bikes with stronger hardware and better firmware, bikes may also communicate with the system. Consequently, to receive the messages of the bike, process them and send a response back, a message broker as well as a service oriented architecture is needed.

Furthermore, a way to save all the information gathered from the use of the system can be implemented using a relational database, as all the information, from bikes, stations and users, like the trip records of each user and the bike's and station's messages, can be inserted into a relational model.

Finally, the users can rent a bike from the many kiosks or using a mobile application, which allows them to view all the available bikes and stations, as well as their trip history. They can also rent a bike, see the progress in real time of their trip and understand when a trip is finished.

The objectives of this dissertation are all related to the development of a machine or deep learning model to predict the number of available bikes in a station at a certain point in time. However, before working on the model it is necessary to understand the data, how it is organized and how it is generated.

The trip data for the model is sourced from an in development BSS. The bikes can communicate with the system which makes the flow of each operation easier. Since the moment a user requests to unlock a bike until the moment the bike is again locked in a station, many messages are exchanged in three different operations:

- **Unlocking a bike**: Whenever a user decides to unlock a bike from a station, the bicycle chosen requests the system to be unlocked, sending a message with the identifier of the user, the bike and the dock of the station. After some verifications the system sends a confirmation or a denial to the bike and the station to either unlock or alert the user, respectively;
- **Riding a bike**: During the trip the bike sends keepalive messages periodically with their state and coordinates, to locate the bike in the system;
- **Locking a bike**: When returning a bike to a station, the bike is automatically locked in the respective dock and the system receives a message from the station informing that the bike is successfully locked.

From these operations many messages are exchanged between the bike and the system, and the station and the system, but the most important messages are the unlocking and locking confirmations of the bike, as it saves the timestamp at when the bike left or arrived at the station and the station the bike left from or arrived to.

CHAPTER 5

Work Plan

This chapter presents an overview of the required tasks, as well as an estimate of the duration of the tasks and the time span of execution.

5.1 Planning

In the first half of this dissertation, a literature review was conducted and some machine and deep learning models were compared for later use, as such the tasks for the next part of this dissertation are listed below:

- Analyze the various prediction models;
  - Start: 13/01/2025
  - End: 07/02/2025
- Choose the best for the objective;
  - Start: 27/01/2025
  - End: 07/02/2025
- Build the model
  - Start: 07/02/2025
  - End: 26/02/2025
- Explore the data to be used and select the best features for our problem
  - Start: 17/02/2025
  - End: 14/03/2025
- Train and test the model, while tuning the hyperparameters for a better result
  - Start: 14/03/2025
  - End: 02/05/2025
- Compare the model created with different state of the art models
  - Start: 07/04/2025
  - End: 02/05/2025
- Validate the model by implementing it in a mobile application for the users of the system

  - Start: 05/05/2025
  - End: 18/06/2025
- Validate the model by implementing it in a reallocation application for the redistributors of the system
  - Start: 05/05/2025
  - End: 18/06/2025

Figure 5.1 shows the tasks and their estimated duration in a Gantt Chart.

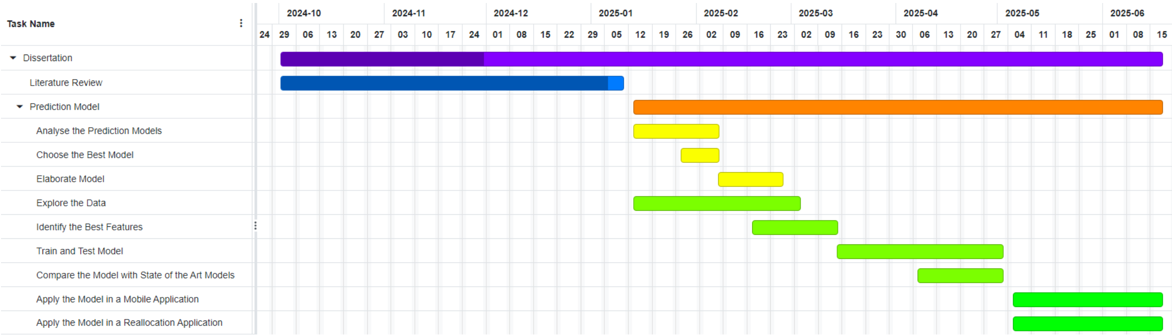

Figure 5.1: Gantt Chart for the Planned Work

CHAPTER 6

Conclusion

Using the Scopus database for scientific publications, an advanced search was conducted that resulted in over 1300 relevant papers, of which 500 papers were used in the evaluation of bike sharing systems. Many scripts were developed to automate and systemize the methodology for the search of documents, which reduced the time spent analyzing the documents.

Five generations of bike sharing systems have been identified, each with greater technologies and solutions to the previous problems, and many benefits, since their origin in 1965. New insights for bike sharing systems were provided, as many bikes were observed parked near train stations and bus stops, indicating a modal shift away from using private cars or taxis.

High demand stations in standard systems and randomly parked bikes in dockless programs, are some of the main problems of these systems. Many studies have been conducted, either to find the most efficient truck routes or to develop and compare the best possible predictive models for the number of available bikes in a station at a certain period of time.

Additionally, the infrastructure, including the database, services, dashboards and applications are very important components that also make up a BSS. The data that will be used to train and test the model is sourced from a system that divides a trip in many operations. Only some information, however, is useful to take from the many messages exchanged between the bike and the system, and between the station and the system, from the start until the end of a trip.

References

[1] S. A. Shaheen, S. Guzman, and H. Zhang, "Bikesharing in europe, the americas, and asia: Past, present, and future," *Transportation Research Record*, vol. 2143, no. 1, pp. 159–167, 2010. DOI: 10.3141/2143-20. eprint: https://doi.org/10.3141/2143-20. [Online]. Available: https://doi.org/10.3141/2143-20.

[2] M. Du and L. Cheng, "Better understanding the characteristics and influential factors of different travel patterns in free-floating bike sharing: Evidence from nanjing, china," *Sustainability*, vol. 10, no. 4, 2018, ISSN: 2071-1050. DOI: 10.3390/su10041244. [Online]. Available: https://www.mdpi.com/2071-1050/10/4/1244.

[3] S. W. Elliot Fishman and N. Haworth, "Bike share: A synthesis of the literature," *Transport Reviews*, vol. 33, no. 2, pp. 148–165, 2013. DOI: 10.1080/01441647.2013.775612. eprint: https://doi.org/10.1080/01441647.2013.775612. [Online]. Available: https://doi.org/10.1080/01441647.2013.775612.

[4] M. Hua, X. Chen, J. Chen, L. Cheng, and D. Lei, "How does dockless bike sharing serve users in nanjing, china? user surveys vs. trip records," *Research in Transportation Business & Management*, vol. 43, p. 100 701, 2022, ISSN: 2210-5395. DOI: https://doi.org/10.1016/j.rtbm.2021.100701. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2210539521000845.

[5] C. Shui and W. Szeto, "Dynamic green bike repositioning problem – a hybrid rolling horizon artificial bee colony algorithm approach," *Transportation Research Part D: Transport and Environment*, vol. 60, pp. 119–136, 2018, Special Issue on Traffic Modeling for Low-Emission Transport, ISSN: 1361-9209. DOI: https://doi.org/10.1016/j.trd.2017.06.023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S136192091730500X.

[6] P. Midgley, "Shared smart bicycle schemes in european cities," *Global Transport Knowledge Partnership (gTKP).(http://www. uncrd. or. jp/env/4th-regional-estforum/Presentations/28_PS4_gTKP. pdf, accessed March 12, 2012)*, 2009.

[7] E. Eren and V. E. Uz, "A review on bike-sharing: The factors affecting bike-sharing demand," *Sustainable Cities and Society*, vol. 54, p. 101 882, 2020, ISSN: 2210-6707. DOI: https://doi.org/10.1016/j.scs.2019.101882. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2210670719312387.

[8] X. Ma, Y. Yuan, N. Van Oort, and S. Hoogendoorn, "Bike-sharing systems' impact on modal shift: A case study in delft, the netherlands," *Journal of Cleaner Production*, vol. 259, p. 120 846, 2020, ISSN: 0959-6526. DOI: https://doi.org/10.1016/j.jclepro.2020.120846. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0959652620308933.

[9] S. A. S. Hua Zhang and X. Chen, "Bicycle evolution in china: From the 1900s to the present," *International Journal of Sustainable Transportation*, vol. 8, no. 5, pp. 317–335, 2014. DOI: 10.1080/15568318.2012.699999. eprint: https://doi.org/10.1080/15568318.2012.699999. [Online]. Available: https://doi.org/10.1080/15568318.2012.699999.

[10] C. News, *Wuhan free rental bikes up to 70,000 intelligent rent but also the system starts*, Retrieved August 9, 2012, 2011. [Online]. Available: http://www.chinanews.com/df/2011/12-31/3575510.shtml.

[11] S. A. Shaheen, H. Zhang, E. Martin, and S. Guzman, "China's hangzhou public bicycle: Understanding early adoption and behavioral response to bikesharing," *Transportation Research Record*, vol. 2247, no. 1, pp. 33–41, 2011. DOI: 10.3141/2247-05. eprint: https://doi.org/10.3141/2247-05. [Online]. Available: https://doi.org/10.3141/2247-05.

[12] R. An, R. Zahnow, D. Pojani, and J. Corcoran, "Weather and cycling in new york: The case of citibike," *Journal of Transport Geography*, vol. 77, pp. 97–112, 2019, ISSN: 0966-6923. DOI: https://doi.org/10.1016/j.jtrangeo.2019.04.016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0966692318307282.

[13] O. Caspi and R. B. Noland, "Bikesharing in philadelphia: Do lower-income areas generate trips?" *Travel Behaviour and Society*, vol. 16, pp. 143–152, 2019, ISSN: 2214-367X. DOI: https://doi.org/10.1016/j.tbs.2019.05.004. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2214367X17300790.

[14] A. Hernandez, M. Raymer, and Y. Chen, "Where did bike-share boom? analyzing impact of infrastructure lockdowns on bike-sharing in chicago," *Transportation Research Interdisciplinary Perspectives*, vol. 23, p. 101 015, 2024, ISSN: 2590-1982. DOI: https://doi.org/10.1016/j.trip.2024.101015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2590198224000010.

[15] H. Yang, Y. Zhang, L. Zhong, X. Zhang, and Z. Ling, "Exploring spatial variation of bike sharing trip production and attraction: A study based on chicago's divvy system," *Applied Geography*, vol. 115, p. 102 130, 2020, ISSN: 0143-6228. DOI: https://doi.org/10.1016/j.apgeog.2019.102130. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0143622819305636.

[16] J. Chibwe, S. Heydari, A. Faghih Imani, and A. Scurtu, "An exploratory analysis of the trend in the demand for the london bike-sharing system: From london olympics to covid-19 pandemic," *Sustainable Cities and Society*, vol. 69, p. 102 871, 2021, ISSN: 2210-6707. DOI: https://doi.org/10.1016/j.scs.2021.102871. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S221067072100161X.

[17] Y. He, Z. Song, Z. Liu, and N. N. Sze, "Factors influencing electric bike share ridership: Analysis of park city, utah," *Transportation Research Record*, vol. 2673, no. 5, pp. 12–22, 2019. DOI: 10.1177/0361198119838981. eprint: https://doi.org/10.1177/0361198119838981. [Online]. Available: https://doi.org/10.1177/0361198119838981.

[18] A. Nikitas, P. Wallgren, and O. Rexfelt, "The paradox of public acceptance of bike sharing in gothenburg," *Proceedings of the Institution of Civil Engineers - Engineering Sustainability*, vol. 169, no. 3, pp. 101–113, 2016. DOI: 10.1680/jensu.14.00070. eprint: https://doi.org/10.1680/jensu.14.00070. [Online]. Available: https://doi.org/10.1680/jensu.14.00070.

[19] S. Narayanan, N. Makarov, E. Magkos, J. M. Salanova Grau, G. Aifadopoulou, and C. Antoniou, "Can bike-sharing reduce car use in alexandroupolis? an exploration through the comparison of discrete choice and machine learning models," *Smart Cities*, vol. 6, no. 3, pp. 1239–1253, 2023, ISSN: 2624-6511. DOI: 10.3390/smartcities6030060. [Online]. Available: https://www.mdpi.com/2624-6511/6/3/60.

[20] T. Fontes, M. Arantes, P. V. Figueiredo, and P. Novais, "A cluster-based approach using smartphone data for bike-sharing docking stations identification: Lisbon case study," *Smart Cities*, vol. 5, no. 1, pp. 251–275, 2022, ISSN: 2624-6511. DOI: 10.3390/smartcities5010016. [Online]. Available: https://www.mdpi.com/2624-6511/5/1/16.

[21] F. Bruzzone, M. Scorrano, and S. Nocera, "The combination of e-bike-sharing and demand-responsive transport systems in rural areas: A case study of velenje," *Research in Transportation Business & Management*, vol. 40, p. 100 570, 2021, Active Travel and Mobility Management, ISSN: 2210-5395. DOI: https://doi.org/10.1016/j.rtbm.2020.100570. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2210539520301085.

[22] F. Dobruszkes and M. Dzięcielski, "Does docked bike-sharing usage complement or overlap public transport? the case of brussels, belgium," *Transportation Planning and Technology*, vol. 47, no. 6, pp. 834–851, 2024. DOI: 10.1080/03081060.2023.2256717. eprint: https://doi.org/10.1080/03081060.2023.2256717. [Online]. Available: https://doi.org/10.1080/03081060.2023.2256717.

[23] B. Laa and G. Emberger, "Bike sharing: Regulatory options for conflicting interests – case study vienna," *Transport Policy*, vol. 98, pp. 148–157, 2020, ISSN: 0967-070X. DOI: https://doi.org/10.1016/j.tranpol.2020.03.009. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0967070X19308881.

[24] S. N. Wahab, R. Sham, A. '. Hussin, S. Ismail, and S. D. Rajendran, "Urban transportation: A case study on bike-sharing usage in klang valley," *International Journal of Supply Chain Management*, vol. 7,

no. 5, pp. 470–476, 2018, Cited by: 14. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85060897447%5C&partnerID=40%5C&md5=6390ae3b8572a76d859e114ab69706c1.

[25] D. v. L. Zheyan Chen and D. Ettema, "Dockless bike-sharing systems: What are the implications?" *Transport Reviews*, vol. 40, no. 3, pp. 333–353, 2020. DOI: 10.1080/01441647.2019.1710306. eprint: https://doi.org/10.1080/01441647.2019.1710306. [Online]. Available: https://doi.org/10.1080/01441647.2019.1710306.

[26] H. Si, J.-g. Shi, G. Wu, J. Chen, and X. Zhao, "Mapping the bike sharing research published from 2010 to 2018: A scientometric review," *Journal of Cleaner Production*, vol. 213, pp. 415–427, 2019, ISSN: 0959-6526. DOI: https://doi.org/10.1016/j.jclepro.2018.12.157. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0959652618338678.

[27] J. Yin, L. Qian, and J. Shen, "From value co-creation to value co-destruction? the case of dockless bike sharing in china," *Transportation Research Part D: Transport and Environment*, vol. 71, pp. 169–185, 2019, The roles of users in low-carbon transport innovations: Electrified, automated, and shared mobility, ISSN: 1361-9209. DOI: https://doi.org/10.1016/j.trd.2018.12.004. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1361920918305595.

[28] M. Hua, X. Chen, S. Zheng, L. Cheng, and J. Chen, "Estimating the parking demand of free-floating bike sharing: A journey-data-based study of nanjing, china," *Journal of Cleaner Production*, vol. 244, p. 118 764, 2020, ISSN: 0959-6526. DOI: https://doi.org/10.1016/j.jclepro.2019.118764. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0959652619336340.

[29] K. Gao, A. Li, Y. Liu, J. Gil, and Y. Bie, "Unraveling the mode substitution of dockless bike-sharing systems and its determinants: A trip level data-driven interpretation," *Sustainable Cities and Society*, vol. 98, p. 104 820, 2023, ISSN: 2210-6707. DOI: https://doi.org/10.1016/j.scs.2023.104820. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2210670723004316.

[30] Y. Fan and S. Zheng, "Dockless bike sharing alleviates road congestion by complementing subway travel: Evidence from beijing," *Cities*, vol. 107, p. 102 895, 2020, ISSN: 0264-2751. DOI: https://doi.org/10.1016/j.cities.2020.102895. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0264275120312439.

[31] X. Ma, R. Cao, and J. Wang, "Effects of psychological factors on modal shift from car to dockless bike sharing: A case study of nanjing, china," *International Journal of Environmental Research and Public Health*, vol. 16, no. 18, 2019, ISSN: 1660-4601. DOI: 10.3390/ijerph16183420. [Online]. Available: https://www.mdpi.com/1660-4601/16/18/3420.

[32] E. Bakogiannis, M. Siti, S. Tsigdinos, A. Vassi, and A. Nikitas, "Monitoring the first dockless bike sharing system in greece: Understanding user perceptions, usage patterns and adoption barriers," *Research in Transportation Business & Management*, vol. 33, p. 100 432, 2019, ISSN: 2210-5395. DOI: https://doi.org/10.1016/j.rtbm.2020.100432. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2210539519302676.

[33] J. Song, L. Zhang, Z. Qin, and M. A. Ramli, "Where are public bikes? the decline of dockless bike-sharing supply in singapore and its resulting impact on ridership activities," *Transportation Research Part A: Policy and Practice*, vol. 146, pp. 72–90, 2021, ISSN: 0965-8564. DOI: https://doi.org/10.1016/j.tra.2021.01.016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0965856421000161.

[34] L. Di Gaspero, A. Rendl, and T. Urli, "Constraint-based approaches for balancing bike sharing systems," in *Principles and Practice of Constraint Programming*, C. Schulte, Ed., Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 758–773, ISBN: 978-3-642-40627-0.

[35] C. Fricker and N. Gast, "Incentives and redistribution in homogeneous bike-sharing systems with stations of finite capacity," *EURO Journal on Transportation and Logistics*, vol. 5, no. 3, pp. 261–291, Aug. 2016, ISSN: 2192-4376. DOI: 10.1007/s13676-014-0053-5.

[36] A. Waserhole and V. Jost, "Vehicle sharing system pricing regulation: Transit optimization of intractable queuing network," *HAL Id: hal-00751744*, pp. 1–20, 2012.

[37] C. Fricker and N. Gast, "Incentives and redistribution in homogeneous bike-sharing systems with stations of finite capacity," *EURO Journal on Transportation and Logistics*, vol. 5, no. 3, pp. 261–291,

2016, ISSN: 2192-4376. DOI: https://doi.org/10.1007/s13676-014-0053-5. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2192437620300959.

[38] B. Legros, "Dynamic repositioning strategy in a bike-sharing system; how to prioritize and how to rebalance a bike station," *European Journal of Operational Research*, vol. 272, no. 2, pp. 740–753, 2019, ISSN: 0377-2217. DOI: https://doi.org/10.1016/j.ejor.2018.06.051. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0377221718306039.

[39] S. Ghosh, P. Varakantham, Y. Adulyasak, and P. Jaillet, "Dynamic repositioning to reduce lost demand in bike sharing systems," *Journal of Artificial Intelligence Research*, vol. 58, pp. 387–430, Feb. 2017, ISSN: 1076-9757. DOI: 10.1613/jair.5308.

[40] L. Caggiani, R. Camporeale, M. Ottomanelli, and W. Y. Szeto, "A modeling framework for the dynamic management of free-floating bike-sharing systems," *Transportation Research Part C: Emerging Technologies*, vol. 87, pp. 159–182, 2018, ISSN: 0968-090X. DOI: https://doi.org/10.1016/j.trc.2018.01.001. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0968090X18300020.

[41] A. Pal and Y. Zhang, "Free-floating bike sharing: Solving real-life large-scale static rebalancing problems," *Transportation Research Part C: Emerging Technologies*, vol. 80, pp. 92–116, 2017, ISSN: 0968-090X. DOI: https://doi.org/10.1016/j.trc.2017.03.016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0968090X17300992.

[42] T. Raviv, M. Tzur, and I. A. Forma, "Static repositioning in a bike-sharing system: Models and solution approaches," *EURO Journal on Transportation and Logistics*, vol. 2, no. 3, pp. 187–229, 2013, ISSN: 2192-4376. DOI: https://doi.org/10.1007/s13676-012-0017-6. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2192437620301175.

[43] M. Dell'Amico, E. Hadjicostantinou, M. Iori, and S. Novellani, "The bike sharing rebalancing problem: Mathematical formulations and benchmark instances," *Omega*, vol. 45, pp. 7–19, 2014, ISSN: 0305-0483. DOI: https://doi.org/10.1016/j.omega.2013.12.001. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0305048313001187.

[44] D. Zhang, C. Yu, J. Desai, H. Lau, and S. Srivathsan, "A time-space network flow approach to dynamic repositioning in bicycle sharing systems," *Transportation Research Part B: Methodological*, vol. 103, pp. 188–207, 2017, Green Urban Transportation, ISSN: 0191-2615. DOI: https://doi.org/10.1016/j.trb.2016.12.006. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0191261516302697.

[45] J. Schuijbroek, R. Hampshire, and W.-J. van Hoeve, "Inventory rebalancing and vehicle routing in bike sharing systems," *European Journal of Operational Research*, vol. 257, no. 3, pp. 992–1004, 2017, ISSN: 0377-2217. DOI: https://doi.org/10.1016/j.ejor.2016.08.029. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0377221716306658.

[46] I. A. Forma, T. Raviv, and M. Tzur, "A 3-step math heuristic for the static repositioning problem in bike-sharing systems," *Transportation Research Part B: Methodological*, vol. 71, pp. 230–247, 2015, ISSN: 0191-2615. DOI: https://doi.org/10.1016/j.trb.2014.10.003. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0191261514001726.

[47] F. Cruz, A. Subramanian, B. P. Bruck, and M. Iori, "A heuristic algorithm for a single vehicle static bike sharing rebalancing problem," *Computers & Operations Research*, vol. 79, pp. 19–33, 2017, ISSN: 0305-0548. DOI: https://doi.org/10.1016/j.cor.2016.09.025. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0305054816302489.

[48] S. C. Ho and W. Szeto, "Solving a static repositioning problem in bike-sharing systems using iterated tabu search," *Transportation Research Part E: Logistics and Transportation Review*, vol. 69, pp. 180–198, 2014, ISSN: 1366-5545. DOI: https://doi.org/10.1016/j.tre.2014.05.017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1366554514000945.

[49] C. Fu, N. Zhu, S. Ma, and R. Liu, "A two-stage robust approach to integrated station location and rebalancing vehicle service design in bike-sharing systems," *European Journal of Operational Research*, vol. 298, no. 3, pp. 915–938, 2022, ISSN: 0377-2217. DOI: https://doi.org/10.1016/j.ejor.2021.06.014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0377221721005294.

[50] L. Lin, Z. He, and S. Peeta, "Predicting station-level hourly demand in a large-scale bike-sharing network: A graph convolutional neural network approach," *Transportation Research Part C: Emerging Technologies*, vol. 97, pp. 258–276, 2018, ISSN: 0968-090X. DOI: https://doi.org/10.1016/j.trc.2018.10.011. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0968090X18300974.

[51] R. Giot and R. Cherrier, "Predicting bikeshare system usage up to one day ahead," in *2014 IEEE Symposium on Computational Intelligence in Vehicles and Transportation Systems (CIVTS)*, 2014, pp. 22–29. DOI: 10.1109/CIVTS.2014.7009473.

[52] X. Zhou, "Understanding spatiotemporal patterns of biking behavior by analyzing massive bike sharing data in chicago," *PLOS ONE*, vol. 10, no. 10, pp. 1–20, Oct. 2015. DOI: 10.1371/journal.pone.0137922. [Online]. Available: https://doi.org/10.1371/journal.pone.0137922.

[53] J. Bao, C. Xu, P. Liu, and W. Wang, "Exploring bikesharing travel patterns and trip purposes using smart card data and online point of interests," *Networks and Spatial Economics*, vol. 17, no. 4, pp. 1231–1253, 2017, Cited by: 127. DOI: 10.1007/s11067-017-9366-x. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85029104015%5C&doi=10.1007%2fs11067-017-9366-x%5C&partnerID=40%5C&md5=1b1d1d23c3798b9716db8c5f2e10109a.

[54] R. Rixey, "Station-level forecasting of bikesharing ridership," *Transportation Research Record*, no. 2387, pp. 46–55, 2013, Cited by: 230. DOI: 10.3141/2387-06. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84897134086%5C&doi=10.3141%2f2387-06%5C&partnerID=40%5C&md5=f31b57b36b793e7553dab9119fae74bd.

[55] A. Faghih-Imani, N. Eluru, A. M. El-Geneidy, M. Rabbat, and U. Haq, "How land-use and urban form impact bicycle flows: Evidence from the bicycle-sharing system (bixi) in montreal," *Journal of Transport Geography*, vol. 41, pp. 306–314, 2014, ISSN: 0966-6923. DOI: https://doi.org/10.1016/j.jtrangeo.2014.01.013. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0966692314000234.

[56] B. Du, X. Hu, L. Sun, J. Liu, Y. Qiao, and W. Lv, "Traffic demand prediction based on dynamic transition convolutional neural network," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 2, pp. 1237–1247, 2021. DOI: 10.1109/TITS.2020.2966498.

[57] C. Xu, J. Ji, and P. Liu, "The station-free sharing bike demand forecasting with a deep learning approach and large-scale datasets," *Transportation Research Part C: Emerging Technologies*, vol. 95, pp. 47–60, 2018, ISSN: 0968-090X. DOI: https://doi.org/10.1016/j.trc.2018.07.013. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0968090X18306764.

[58] S.-H. Lee and H.-C. Ku, "A dual attention-based recurrent neural network for short-term bike sharing usage demand prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 4, pp. 4621–4630, 2023. DOI: 10.1109/TITS.2022.3208087.

[59] G. Xiao, R. Wang, C. Zhang, and A. Ni, "Demand prediction for a public bike sharing program based on spatio-temporal graph convolutional networks," *Multimedia Tools and Applications*, vol. 80, no. 15, pp. 22 907–22 925, Mar. 2020, ISSN: 1573-7721. DOI: 10.1007/s11042-020-08803-y.

[60] X. Li, Y. Xu, X. Zhang, W. Shi, Y. Yue, and Q. Li, "Improving short-term bike sharing demand forecast through an irregular convolutional neural network," *Transportation Research Part C: Emerging Technologies*, vol. 147, p. 103 984, 2023, ISSN: 0968-090X. DOI: https://doi.org/10.1016/j.trc.2022.103984. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0968090X22003977.

[61] A. Mehdizadeh Dastjerdi and C. Morency, "Bike-sharing demand prediction at community level under covid-19 using deep learning," *Sensors*, vol. 22, no. 3, 2022, ISSN: 1424-8220. DOI: 10.3390/s22031060. [Online]. Available: https://www.mdpi.com/1424-8220/22/3/1060.

[62] X. Ma, Y. Yin, Y. Jin, M. He, and M. Zhu, "Short-term prediction of bike-sharing demand using multi-source data: A spatial-temporal graph attentional lstm approach," *Applied Sciences*, vol. 12, no. 3, 2022, ISSN: 2076-3417. DOI: 10.3390/app12031161. [Online]. Available: https://www.mdpi.com/2076-3417/12/3/1161.

[63] Y. Yang, A. Heppenstall, A. Turner, and A. Comber, "Using graph structural information about flows to enhance short-term demand prediction in bike-sharing systems," *Computers, Environment and Urban Systems*, vol. 83, p. 101 521, 2020, ISSN: 0198-9715. DOI: https://doi.org/10.1016/

j.compenvurbsys.2020.101521. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0198971520302544.

[64] W. Zi, W. Xiong, H. Chen, and L. Chen, "Tagcn: Station-level demand prediction for bike-sharing system via a temporal attention graph convolution network," *Information Sciences*, vol. 561, pp. 274–285, 2021, ISSN: 0020-0255. DOI: https://doi.org/10.1016/j.ins.2021.01.065. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0020025521001031.

[65] M. Van Der Voort, M. Dougherty, and S. Watson, "Combining kohonen maps with arima time series models to forecast traffic flow," *Transportation Research Part C: Emerging Technologies*, vol. 4, no. 5, pp. 307–318, 1996, Cited by: 719; All Open Access, Green Open Access. DOI: 10.1016/S0968-090X(97)82903-8. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-00302989511%5C&doi=10.1016%2fS0968-090X%2897%2982903-8%5C&partnerID=40%5C&md5=3ae807e44949e76b5c6fe8ed875d1543.

[66] W. Jia, Y. Tan, L. Liu, J. Li, H. Zhang, and K. Zhao, "Hierarchical prediction based on two-level gaussian mixture model clustering for bike-sharing system," *Knowledge-Based Systems*, vol. 178, pp. 84–97, 2019, ISSN: 0950-7051. DOI: https://doi.org/10.1016/j.knosys.2019.04.020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0950705119301935.

[67] A. Faghih-Imani, R. Hampshire, L. Marla, and N. Eluru, "An empirical analysis of bike sharing usage and rebalancing: Evidence from barcelona and seville," *Transportation Research Part A: Policy and Practice*, vol. 97, pp. 177–191, 2017, ISSN: 0965-8564. DOI: https://doi.org/10.1016/j.tra.2016.12.007. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0965856416311648.

[68] A. Kaltenbrunner, R. Meza, J. Grivolla, J. Codina, and R. Banchs, "Urban cycles and mobility patterns: Exploring and predicting trends in a bicycle-based public transport system," *Pervasive and Mobile Computing*, vol. 6, no. 4, pp. 455–466, 2010.

[69] Y. Li and Y. Zheng, "Citywide bike usage prediction in a bike-sharing system," *IEEE Transactions on Knowledge and Data Engineering*, vol. 32, no. 6, pp. 1079–1091, 2020. DOI: 10.1109/TKDE.2019.2898831.

[70] W. Jiang, "Bike sharing usage prediction with deep learning: A survey," *Neural Computing and Applications*, vol. 34, no. 18, pp. 15 369–15 385, Jun. 2022, ISSN: 1433-3058. DOI: 10.1007/s00521-022-07380-5.

[71] H. I. Ashqar, M. Elhenawy, and H. A. Rakha, "Modeling bike counts in a bike-sharing system considering the effect of weather conditions," *Case Studies on Transport Policy*, vol. 7, no. 2, pp. 261–268, 2019, ISSN: 2213-624X. DOI: https://doi.org/10.1016/j.cstp.2019.02.011. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2213624X16301018.

[72] Y. Lv, D. Zhi, H. Sun, and G. Qi, "Mobility pattern recognition based prediction for the subway station related bike-sharing trips," *Transportation Research Part C: Emerging Technologies*, vol. 133, p. 103 404, 2021, ISSN: 0968-090X. DOI: https://doi.org/10.1016/j.trc.2021.103404. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0968090X21004009.

[73] S. V. E. and Y. Cho, "Season wise bike sharing demand analysis using random forest algorithm," *Computational Intelligence*, vol. 40, no. 1, e12287, 2024. DOI: https://doi.org/10.1111/coin.12287. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/coin.12287. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/coin.12287.

[74] S. V. E, J. Park, and Y. Cho, "Using data mining techniques for bike sharing demand prediction in metropolitan city," *Computer Communications*, vol. 153, pp. 353–366, 2020, ISSN: 0140-3664. DOI: https://doi.org/10.1016/j.comcom.2020.02.007. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0140366419318997.

[75] X. Wang, Z. Cheng, M. Trépanier, and L. Sun, "Modeling bike-sharing demand using a regression model with spatially varying coefficients," *Journal of Transport Geography*, vol. 93, p. 103 059, 2021, ISSN: 0966-6923. DOI: https://doi.org/10.1016/j.jtrangeo.2021.103059. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0966692321001125.

[76] T. Xu, G. Han, X. Qi, J. Du, C. Lin, and L. Shu, "A hybrid machine learning model for demand prediction of edge-computing-based bike-sharing system using internet of things," *IEEE Internet of Things Journal*, vol. 7, no. 8, pp. 7345–7356, 2020. DOI: 10.1109/JIOT.2020.2983089.

[77] S. V E and Y. Cho, “A rule-based model for seoul bike sharing demand prediction using weather data,” *European Journal of Remote Sensing*, vol. 53, no. sup1, pp. 166–183, 2020, Cited by: 57; All Open Access, Gold Open Access. DOI: 10.1080/22797254.2020.1725789. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85079721316%5C&doi=10.1080%2f22797254.2020.1725789%5C&partnerID=40%5C&md5=2d9f1f30c23dcef31afc7cb8cb9603dc.