Title : Multi-scale prediction of bike-share availability from open data

Scientific domains : computer science, mathematics, geomatics

Key-Words : bike sharing, open data, time series analysis

# Supervision

## Supervisor (s) :

*Gaël PALLARES, Lecturer-researcher*

*Myriam FOUCRAS, Lecturer-researcher*

# Research Work

## Abstract of the internship

In the context of climate change and dwindling natural resources, it is becoming vital to optimize existing transport systems and develop the use of more ecologically « virtuous » modes, such as bike-sharing. The design and management of a transport network requires detailed knowledge of demand at all scales. From the user's point of view, high demand may mean that shared transport is unavailable at a given moment, with the result that the mode may be rejected. In this context, it is crucial to be able to deliver this information in real time to future passengers. The deployment of digital tools enables the automatic acquisition of data at any given moment, and their analysis appears to be a promising way to provide such information to users. The aim of this internship is to use open data to predict the availability of bike-share at stations on different timescales.

## Internship Project

### Scientific context

This internship is part of CESI's MobE program. This program aims to understand and optimize student mobility in favor of energy savings, and to engage higher education establishments in a transformation of student mobility via an integrated approach of analysis, methodological input and wide dissemination to students and higher education establishments. Our aim is to reduce the carbon footprint of student travel, while proposing efficient alternatives based on soft mobility (walking, cycling, scootering, etc.) and any other means of mobility, whether collective or individual, that contribute to a reduction in CO2 emissions (such as public transport or car-sharing).

### Internship subject

The densification of urban populations is leading to an increase in the need for urban mobility, and with it the aggravation of various underlying problems, notably energy and ecological. In a context of finite space and resources, it is vital to optimize existing transport systems, and to develop the use of more ecologically « virtuous » modes (Gallo and Marinelli, 2020). Bike-sharing is one of these sustainable modes of transport. This public mobility service enables users to borrow and return bicycles from stations strategically located around the city. To guarantee the fluidity and accessibility of the system, it is essential to ensure that bikes and parking spaces are available at the right time and in the right place: users need to find an available bike on departure and a free parking space on arrival. However, this availability is variable and can depend on many factors, such as weather conditions, time of day, and people's travel habits.

Several modeling approaches are possible to meet the objectives of this project. The first is time-series prediction. By studying temporal variations in bicycle availability, it is possible to detect recurring

patterns according to time of day, day of the week and/or season (Kaltenbrunner et al., 2010). Other machine learning approaches are also possible: advanced prediction models such as random forests, regression models or recurrent neural networks can potentially be employed to capture complex relationships between the different variables (Ashqar et al., 2017).

### Prior art study in the laboratory

Several doctoral research are funded by the « Mon Trajet Vert » program. One thesis in particular deals with mobility planning for higher education students, using a dynamic multimodal system. This internship will contribute to this thesis by providing data and tools for predicting the availability of the soft mode. Unlike public transport or trains, which are scheduled, no predictive data is provided for shared bikes.

### Work program

| Périodes | Etapes |
|---|---|
| **Month 1** | Bibliography and action plan definition |
| **Month 2-3** | Data cleaning and pre-analysis<br>Prediction model development |
| **Month 3-4** | Prediction model deployment on the real dataset<br>Evaluation of metrics and discussion of model limitations |
| **Month 6** | Visualizations and written (report) and oral presentations |

### Expected scientific/technical production

The expected deliverables are :
- the internship report including the prediction model explained
- the associated Python programming code
- the data obtained (spatio-temporal availability).

This work will provide an additional « building block » for understanding the levers and barriers to the bike use, particularly among students (Torrisi et al, 2021). It can also be used as input data for simulation tools designed to describe and predict the dynamics of mobility flows (Hörl and Balac, 2021). It will also be useful as dataset for the study of bicycle fleet rebalancing optimization problems in operations research (Schuijbroek et al., 2017) and will enable a description of travel by weight assignment in graph theory (Yang et al., 2020). Depending on the progress of the internship and the results obtained, a scientific production is envisaged.

## Context

### Lab presentation

CESI LINEACT (UR 7527), Laboratory for Digital Innovation for Businesses and Learning to Support the Competitiveness of Territories, anticipates and accompanies the technological mutations of sectors and services related to industry and construction. The historical proximity of CESI with companies is a determining element for our research activities, and has led us to concentrate our efforts on applied research close to the company and in partnership with them. A human-centered approach coupled with the use of technologies, as well as territorial networking and links with training, have enabled the construction of cross-cutting research; it puts humans, their needs and their uses, at the center of its issues and addresses the technological angle through these contributions.
Its research is organized according to two interdisciplinary scientific teams and several application areas.

- Team 1 "Learning and Innovating" mainly concerns Cognitive Sciences, Social Sciences and Management Sciences, Training Techniques and those of Innovation. The main scientific objectives are the understanding of the effects of the environment, and more particularly of situations instrumented by technical objects (platforms, prototyping workshops, immersive systems...) on learning, creativity and innovation processes.
- Team 2 "Engineering and Digital Tools" mainly concerns Digital Sciences and Engineering. The main scientific objectives focus on modeling, simulation, optimization and data analysis of cyber physical systems. Research work also focuses on decision support tools and on the study of human-system interactions in particular through digital twins coupled with virtual or augmented environments.

These two teams develop and cross their research in application areas such as

- Industry 5.0,
- Construction 4.0 and City of the Future,
- Digital Services.

Areas supported by research platforms, mainly that in Rouen dedicated to Factory 5.0 and those in Nanterre dedicated to Factory 5.0 and Construction 4.0.

### Links to the research themes of the research team involved.

The proposed internship deals with issues of shared mobility, and its aim is essentially in an applied research scope. This work is linked with previous research activities of the LINEACT laboratory - Multimodal Transport Systems Management research axis, sub-axis: Shared Mobility of CESI LINEACT Research Theme 2 - Engineering and Digital Tools.

## Internship organisation

Funding : MTV program
Workplace : CESI Toulouse or CESI Montpellier
Starting date : February 2024
Duration of the internship : 6 months

## Recruitment terms and conditions

Terms : by application and interview.
Please send your application to gpallares@cesi.fr with the following subject:
**"[Application] Intership_BikePred".**
Your application must include :

- A detailed Curriculum-Vitae. In the event of a break in the academic curriculum, please give an explanation;
- A cover letter explaining why you wish to pursue an internship;
- MASTER 1 and 2 results (to be adapted to the level of the internship) and corresponding transcripts;
- any other documents you consider useful.

**Please send all documents in a zip file entitled NAME firstname.zip.**

## Skills :

*Scientific and technical skills :*

- Mathematical modeling,
- Data processing skills,
- Python programming.

*Relational skills :*

- Autonomy, initiative and curiosity,
- Ability to work as part of a team, and good interpersonal skills,
- Rigorous.

# Bibliography.


Ashqar, H. I., Elhenawy, M., Almannaa, M. H., Ghanem, A., Rakha, H. A., & House, L. (2017, June). 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)* (pp. 374-378). IEEE.

Gallo, M., & Marinelli, M. (2020). Sustainable mobility: A review of possible actions and policies. *Sustainability*, *12*(18), 7499.

Hörl, S., & Balac, M. (2021). Synthetic population and travel demand for Paris and Île-de-France based on open and publicly available data. *Transportation Research Part C: Emerging Technologies*, *130*, 103291.

Kaltenbrunner, A., Meza, R., Grivolla, J., Codina, J., & Banchs, R. (2010). Urban cycles and mobility patterns: Exploring and predicting trends in a bicycle-based public transport system. *Pervasive and Mobile Computing*, *6*(4), 455-466.

Schuijbroek, J., Hampshire, R. C., & Van Hoeve, W. J. (2017). Inventory rebalancing and vehicle routing in bike sharing systems. *European Journal of Operational Research*, *257*(3), 992-1004.

Torrisi, V., Ignaccolo, M., Inturri, G., Tesoriere, G., & Campisi, T. (2021). Exploring the factors affecting bike-sharing demand: Evidence from student perceptions, usage patterns and adoption barriers. *Transportation Research Procedia*, *52*, 573-580.

Yang, Y., Heppenstall, A., Turner, A., & Comber, A. (2020). Using graph structural information about flows to enhance short-term demand prediction in bike-sharing systems. *Computers, Environment and Urban Systems*, *83*, 101521.