Università degli studi di Milano-Bicocca

Data Science Lab for Smart Cities

Final Essay

E-scooters drop-off forecasting: a light LSTM approach

Authors:
Remo Marconzini - 883256 - r.marconzini@campus.unimib.it
Niccolò Puccinelli - 881395- n.puccinelli@campus.unimib.it

June 11, 2023

Abstract

In recent years, the advent and rapid spread of e-scooters has provoked increasing attention from researchers, who have studied the phenomenon from several perspectives and proposed many policies as a result of their findings. In this paper we propose a different approach to the problem of e-scooter availability, more focused on the end user. Firstly, using a week's worth of data (Dec. 2, 2019 to Dec. 8, 2019) of e-scooters in the city of Milan, through a simple recurrent neural network with LSTM units, we aim to predict, given a sequence of e-scooters with the same destination, their common drop-off zone. Next, the same task is performed with an even greater level of granularity, considering the drop-off zone of each individual e-scooter. The study was conducted by considering two distinct numbers of e-scooters in each sequence (5 and 10) and two different levels of granularity of the city, performing a classification task by dividing the city of Milan into 12 and 36 discrete zones, respectively. This simple LSTM model can bring several benefits, both to operators and public authorities and end users. By forecasting trips in different areas of Milan, operators and public authorities would have more control over their fleet in terms of relocation, recharging and managing any peaks and drops. On the other hand, users far from the first available e-scooter would be able to monitor the likelihood that one of these vehicles will be left in their target area.

1 Introduction

The smart city concept has proven to be variable and flexible over the years. According to [1], starting from a mostly technological approach, governance and the theme of sustainability have gradually gained more importance. The latest, post-pandemic phase, has placed greater emphasis on community, surveillance, and public-private partnerships. At present, the different dimensions of the smart city are combined within two main concepts: technology and society (i.e., how and for whom should the city be?). The interaction between these two main elements, together with the work of public and private actors, shapes the smart city. Therefore, in order to broaden the scope of definition of the smart city, we may state that its four basic characteristics are: technology, infrastructure, governance, and people.

Despite the lack of a precise reference architecture, two main contrasting approaches can be identified: technology-oriented (or top-down) and people-oriented. In the former, citizens are involved to a very limited extent, policy decisions are exclusively top-down, while the flow of data, which is also unidirectional, is bottom-up. In the second approach, city data are open to citizens, who are encouraged to participate in decision-making processes. From the perspective of integrating the main features of the smart city, the second approach is clearly preferable, in order to overcome the new urban challenges in the city context. Indeed, direct citizen involvement can substantially contribute to urban planning, improving sustainability (environmental, civic, and economic), quality of life for citizens, and resilience, i.e. "the ability to prevent actual or potentially adverse events from occurring, to take them, cope with them and adapt to them more and more successfully" [2].

Within this context, various communities around the world have shown their interest in smart cities through multiple initiatives. In the European sphere, a noteworthy initiative called the European Smart Cities 4.0 [3] has emerged. This project aims to assess and benchmark the level of smartness in various European medium-sized cities (i.e. from 300.000 to 1 million inhabitants), across diverse domains such as transportation, waste management, and energy. The use of such benchmarking indicators is a clear example of urban informatics, the technological foundation of a smart city undertaking, described by Foth et al. [4] as the encounter and interaction between computer science, people and urban spaces. Within the realm of smart cities, urban informatics finds diverse applications encompassing energy management, healthcare, security, governance, and economics.

Our research endeavors take place within this multifaceted and intricate landscape. Specifically, we decided to consider the topic of smart urban mobility, focusing on the utilization of data obtained from various e-scooters deployed across the city of Milan. The primary objective of this paper is to establish a comprehensive framework that addresses the lack of e-scooters from the perspective of the user. Instead of suggesting relocation policies, we conducted an analysis of the e-scooters time series data and developed a lightweight neural network capable of predicting the likely drop-off point of an e-scooter. Consequently, in the event of vehicle unavailability, prospective users can check the status of each e-scooter and its attached probabilities that it will stop in a nearby area.

Furthermore, the patterns uncovered through data exploration and the application of the neural network have the potential to benefit other stakeholders, including operators, service providers, and public entities. By acquiring a deeper understanding of e-scooters patterns, these actors can enhance their capacity to effectively manage and organize the flow of e-scooters with a heightened sense of awareness.

To assess its robustness and generalization capabilities, we employed an identical network structure for both research questions, progressively increasing in complexity:

1. Can the same drop-off zone for a sequence of e-scooters be predicted by analyzing their individual trajectories?
2. Can the drop-off zone of each individual e-scooter be predicted by analyzing its individual trajectory?

The work is structured as follows: in Section 2 we provide a brief review of relevant literature pertaining to the subject. Section 3 introduces the concept of smart mobility with a particular focus on Milan and micro-mobility, encompassing e-scooters. Following the dataset description in Section 4, we proceed to outline the problem statement in Section 5. Specifically, we investigate whether there are specific times and areas in Milan that experience an excess or shortage of e-scooters. We describe the analysis and prediction methodology in Section 6, while the the neural network results are discussed in Section 7. Section 8 presents a summary of our work and final conclusions, along with possible future developments and enhancements.

2 Related work

The outbreak of the COVID-19 pandemic has significantly impacted various aspects of our daily routines, including transportation. Among the segments profoundly affected, Local Public Transport (LPT) stands out due to the heightened risk of infection. A study [5] conducted among in the cities of Naples and Benevento on university students, one of the categories most involved in the use of public transportation, revealed the drastic shift in their habits after the pandemic. Subsequently, there has been a notable rise in the use of e-scooters by university students, but they're still not considered as valid substitutes for the private car.

Also in the context of Italian cities, considering key factors such as safety, transport requirements, and land use characteristics, the authors of [6] proposed a mobility network plan for e-scooters in Catania, a medium-sized city that suffers from a lack of adequate infrastructure for sustainable transport modes.

As for the city of Milan, in 2018 *Assolombarda*, an association of Lombard industrialists, began promoting a strategic plan, including other companies, to improve Milan smart factor through projects for concrete interventions in 9 sectors (including mobility) and 4 infrastructure layers. Originally known as *Smart-Milano 2030*, in 2021 this public-private initiative became the *Milano Smart City Alliance* [7], which is currently comprised of notable companies (e.g. *Accenture*, *A2A*, *IBM*, *FastWeb*) and is more focused toward issues such as sustainability, resilience, and mobility.

In 2021, Milan and e-scooters were the subject of a research [8] focused on the number of road accidents, which proved the necessity to better plan and design transport infrastructures, while, more recently, authors of [9] made an exhaustive and comprehensive review of the actual state of smart mobility in Milan using a data-driven statistical approach.

The same theme also assumes relevance when comparing Milan with other European cities.

In 2020, a joint working group of *Assolombarda* and *EY* members produced a booklet [10] comparing Milan with four other European cities considered similar in terms of role and economic vocation: Barcelona, Munich, Lyon and Stuttgart. All of these realities share three cross-cutting directions for their future: the smart city as a lever for marketing and attractiveness, the push for alternative mobility and the strategic use of data. The framework analyzes data available as of November 2019 and is structured into 4 integrated layers (Infrastructure and Networks, Sensors, Service Delivery Platform, Applications and Services) and 134 total indicators.

The authors of [11] conducted a comparative analysis of micro-mobility usage across 30 distinct European cities, exploring the temporal and statistical distribution of trips and unveiling both similarities and differences in usage patterns among these cities. The primary objective of this investigation was to enhance efficiency by minimizing electricity waste, thereby optimizing the overall performance of these systems in urban environments.

A comprehensive analysis [12] on the usage and integration of e-scooters has also been conducted in Munich, one of the previously mentioned cities similar to Milan. The findings of this research indicate that e-scooters are perceived more as a leisure/fun object rather than an alternative to private transportation. Interestingly, e-scooters mainly replace other environmentally friendly modes of mobility such as walking and TPL, while the impact on private car usage remains minimal.

The topic has also been extensively addressed from a more technical point of view.

The authors of [13] proposed an e-scooter demand model capable of simulating diverse fleet management policies, with a focus on battery charging strategies. The model benefits from the availability of extensive open data, sourced from the municipalities of Minneapolis and Louisville.

The prediction of the demand has also been addressed by [14], employing deep learning techniques to develop a rebalancing strategy that anticipates the demand patterns of e-scooters. This research utilized one month of data sourced from a South Korean company, enabling the authors to train and evaluate their predictive models.

Relocation is another critical aspect frequently tackled in the realm of e-scooters. A notable study by the authors of [15] highlighted that efficient relocation policies can significantly enhance demand satisfaction by 42%. The study was conducted on the cities of Austin and Louisville, using a simulative approach driven by deep learning techniques.

Lastly, [16], dealt with the next position prediction of moving objects. Leveraging real-world data, this study addressed the task by employing artificial neural networks (ANNs) with LSTM layers, achieving state-of-the-art performances.

To the best of our knowledge, this research, concerning e-scooters drop-off forecasting in the city of Milan through a light LSTM network, represents the first contribution in this specific field.

3 Smart mobility: an overview

The smartness of a city is largely realized through the meticulous study and strategic planning of urban mobility. Smart mobility, a focal area within the realm of urban informatics, revolves around establishing an efficient and effective mobility system by leveraging innovative technologies consistently and systematically. The construction of new long-term policies should go through an "avoid-shift-improve" approach [17], that is, avoid the need for mobility, aim for a shift to sustainable means of transportation and improve existing components of the system. The ultimate objective is to connect the city resources, encompassing people, goods and information.

In recent years, a new mobility paradigm has emerged: Mobility as a Service (MaaS), a holistic approach to mobility that aims to offer a wide range of transportation options as an integrated service. The integration of this novel solution involves the cooperation of a multitude of actors across different levels [18]:

- Public and regulatory level, which includes both national and local authorities.
- Service providers level, e.g. car-sharing companies.
- MaaS operator level, whose role is to combine transportation services into a single mobile application.
- End user level.

Nevertheless, the MaaS model is not exempt from encountering certain challenges. One prominent issue arises from the misalignment of objectives among different actors involved. Public authorities are primarily concerned with promoting sustainability and addressing broader societal needs, whereas private actors are predominantly driven by market forces and profit-oriented goals. Therefore, it is necessary a collective effort to reconcile and harmonize the diverse interests and priorities, in order to ensure the inclusiveness and sustainability of the service. Second, the massive use of this new approach to mobility could lead to an increase in car trips, even among those who would typically choose walking or use public transportation. As previously pointed out in [5] and [12], the introduction of scooter sharing in Munich and Naples (aided by the COVID-19 pandemic), led to a decrease in the utilization of public transportation rather than private cars. Moreover, reports on mobility systems in Milan [19] have revealed a discernible decline in subway usage alongside a significant surge in the popularity of shared mobility services. Lastly, a MaaS transportation system could lead to further inequality on multiple fronts: economically, spatially (with vehicles concentrated in already well-served areas) and socially, where the digital divide becomes a crucial element.

Looking more specifically at sharing mobility, it primarily encompasses two macro-categories: cars and micro-mobility. Micro-mobility, consisting of scooters, bicycles, and electric kick scooters, has witnessed a remarkable surge in recent years. As highlighted by [5][8][9][12], the main characteristics of micro-mobility are:

- The prevalent use of young people and university students.
- The very short commutes.
- The predominant application of these services as entertainment and not as an alternative to private cars.
- Various common technical issues, mainly related to batteries.
- Security issues.

Similar to car-sharing, micro-mobility can be categorized into two main usage patterns: station-based (with fixed pick-up and drop-off points) and free-flow, with the latter posing greater challenges related to relocation.

3.1 The case of Milan

Milan, "the Italian Smart City"[9], is the first Italian city in terms of sharing mobility, according to the 2022 National Sharing Mobility Report [20], wherein sharing services analyzed include cars, bicycles, scooters and electric kick scooters. Although Milan is not yet fully integrated with the MaaS model (the mobility scenario is still very fragmented and lacks interoperability), some commonalities can be seen with the strengths and weaknesses of the new paradigm. Furthermore, despite its advancements (even compared to several European contexts), Milan still faces challenges in the development of an urban data representation platform, as underscored by [10]. While there has been notable progress in the publication of datasets on the city open data portal, there remains a scarcity of comprehensive and complete data specifically related to sharing mobility. This data gap hinders the analysis and understanding of sharing mobility dynamics within the city.

The most recent data available for the vehicle fleet of Milan residents is from 2019 [21], distributed as in fig. 1.

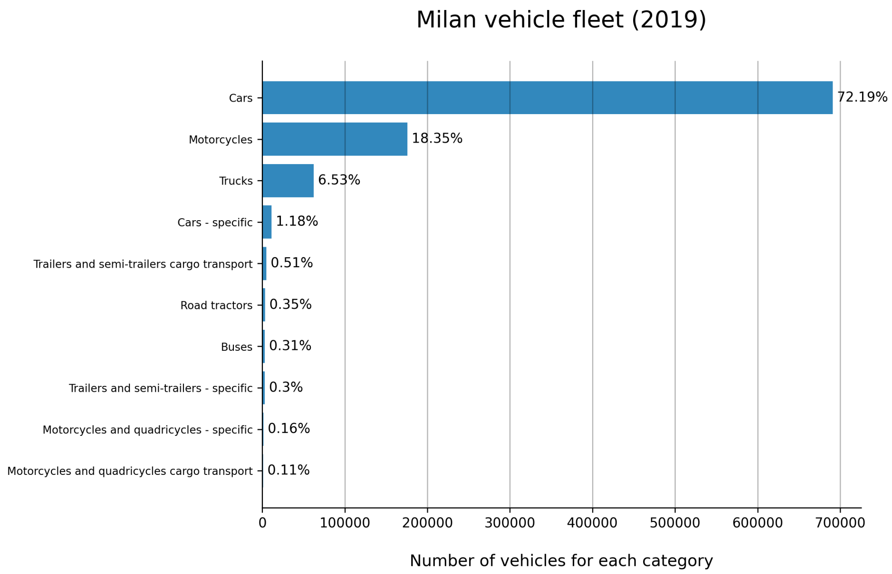


Figure 1: Vehicular fleet of Milan residents - 2019

Although the majority of citizens prefer environmentally friendly travel (i.e. public transport, walking, cycling), 2021 data [22] still show that more than 30% of people predominantly use private cars (see fig. 2).

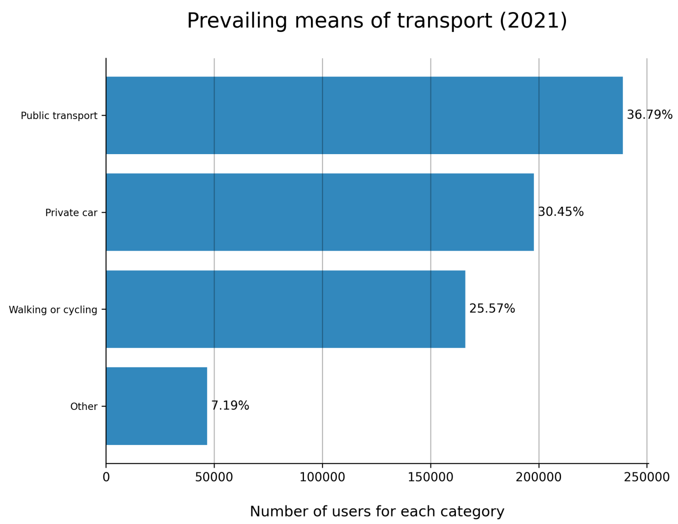


Figure 2: Means of transportation preferred by the citizens of Milan - 2021

Interestingly, these data are grouped by NIL (i.e. *Nuclei di Identità Locale*), which are different areas that can be defined as neighborhoods of Milan. By merging these information with public transport data, we can discover additional insights. In particular, we combined the route of the metro lines with the preferred means of citizens in each NIL. Figure 3 shows how the preferences are mainly influenced by two factors: the proximity to the center of the city and the proximity to a subway stop. This demonstrates that, through an adequate public transportation service, the majority of citizens are willing to abandon the use of the private car. In addition, the introduction of "Area C" in 2012 imposed access restrictions on various categories of polluting vehicles in the city center. Over the years, the restrictions on gasoline and diesel vehicles have become more and more stringent. This measure has significantly impacted the adoption of electric solutions, such as shared mobility vehicles, and has further encouraged the use of public transportation, as well as walking and cycling.

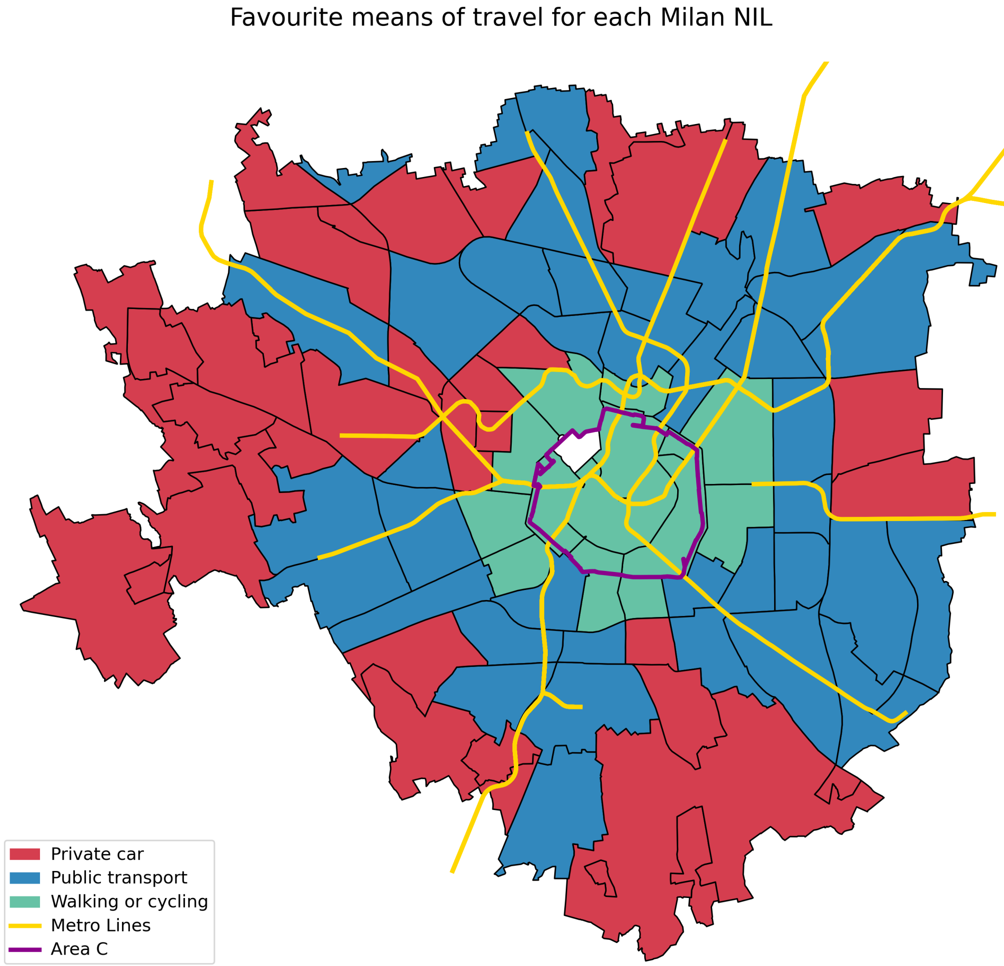


Figure 3: Means of transportation preferred by the citizens of Milan for each NIL - 2021

According to [20], it is noteworthy that by 2021, a total of 62 Italian cities had embraced sharing services, with multiple operators actively participating. Milan and Rome emerged as the leading cities in this domain. From a sustainability perspective, much progress has been made: 94.5% of the sharing fleet in Italy was zero-emission in 2021. As for Milan, the open data platform offers a dataset [23] regarding the general trend of sharing mobility from 2011 to 2021. The treemap in figure 4 shows that, in 2021, the most used service turns out to be bike-sharing (58.8%), while car-sharing (7.7%) ranks below electric kick scooters (18.2%) and scooter-sharing (15.3%). Among rental modes, the free-flow option clearly prevails, constituting 80.8% of the total usage.

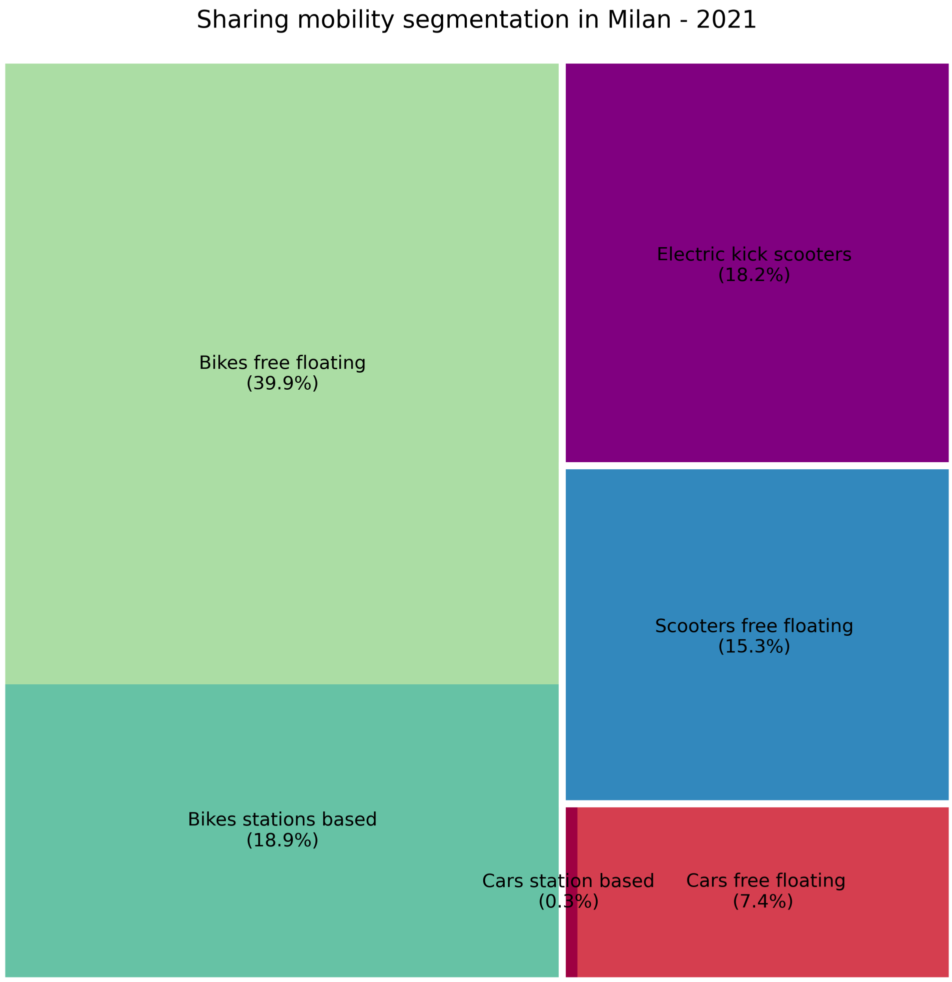
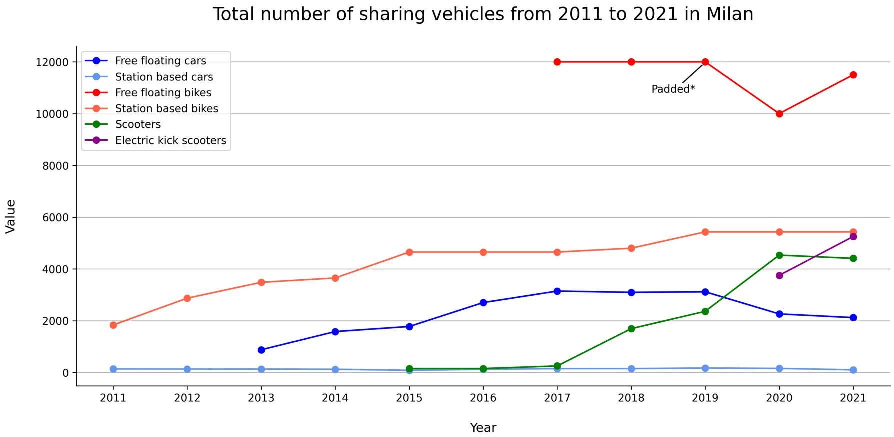


Figure 4: Sharing mobility segmentation in Milan - 2021

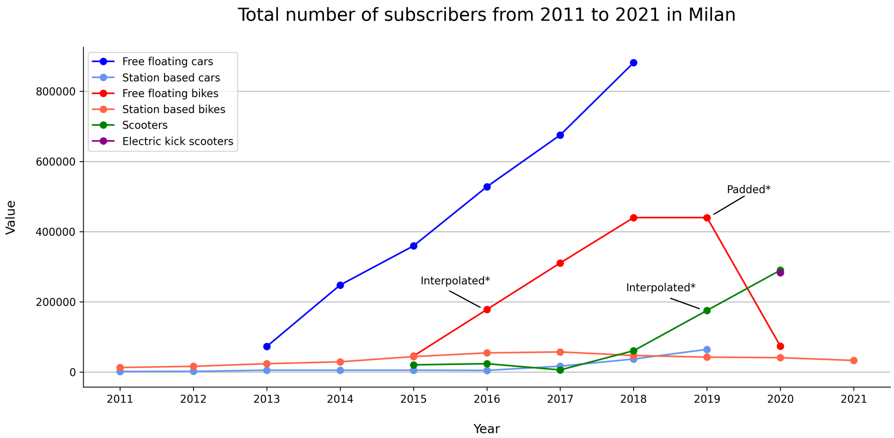
The 2 line plots below (fig. 5) represent the annual trends in the number of vehicles and subscriptions for the previously mentioned sharing services in the city of Milan. There are some missing values in the dataset (mainly related to 2019) that have been inferred. Specifically, we opted to apply padding with the last available value if the trend was ascending before 2020 and experienced a decline thereafter. For cases where the trend did not exhibit such behavior, we chose to interpolate the missing values. Furthermore, if there were more than one missing data point, no imputation was performed.

The first services to emerge were station-based cars and bikes, in 2011. In 2013 came the first free-flowing service, related to car-sharing. General growth can be seen for all categories, with a decline in 2020 due to COVID-19. Scooters show a high growth rate from 2017 to 2020, and in general, despite coming later, free-floating vehicles are predominant. Regarding the number of subscribers, the dataset is largely incomplete (e.g. data on free-floating car-sharing from 2019 onward are missing), but it is interesting to remark the growth regarding scooters contrasted with the drastic decline in free-flowing bikes.

In summary, once again the COVID-19 pandemic turns out to be crucial with regard to urban mobility. Lockdown and the fear of being infected have had a great influence on sharing mobility as well as public transportation. Updated graphs provided by [19], along with several previously cited articles, show a shift in citizens habits toward alternative forms of mobility which are perceived as providing greater safety compared to traditional public transportation methods.



(a) Yearly time series of the number of sharing vehicles in Milan - from 2011 to 2021



(b) Yearly time series of the number of subscribers to sharing vehicles in Milan - from 2011 to 2021

Figure 5: Yearly time series of sharing vehicles and subscribers in Milan

4 Dataset description

The dataset at our disposal pertained to the trajectories of diverse e-scooters within Milan city, spanning from 2019-12-02 00:10:10 to 2019-12-08 23:55:30. There are a total of 82687 samples and 19 columns:

- *provider_id*: ID assigned to the specific e-scooter.
- *lat*: Latitude of the observation.
- *lng*: Longitude of the observation.
- *lat&lng*: Tuple of the geographic coordinates of the observation.
- *created_at*: Datetime of the observation recording.
- *placename*: Marker ID that allows the identification of a city area, i.e. a discretization of geographic coordinates. There are a total of 224 different ID markers.
- *placelat*: Latitude of the marker ID.
- *placelon*: Longitude of the marker ID.
- *distance*: Geographic distance between the marker and the effective coordinates.
- *prevObservation*: Tuple of the geographic coordinates of the previous observation.
- *prevCreated_at*: Datetime of the previous observation.

- *prevDistance*: Previous geographic distance between the marker and the effective coordinates.
- *prevTimeDiff*: Time difference between previous and current observation.
- *nextObservation*: Tuple of the geographic coordinates of the next observation.
- *nextCreated_at*: Datetime of the next observation.
- *nextDistance*: Geographic distance between the marker and the effective coordinates.
- *nextTimeDiff*: Time difference between current and next observation.
- *samePlace*: Binary variable that is worth 1 if the position of the e-scooter is the same as the previous observation.
- *type*: Vehicle status, can take 7 different categorical values:
  - *pick-up*: The scooter was taken.
  - *pick-up outlier*: The scooter pick-up point is far from any placename ID.
  - *first seen / pick-up*: The scooter was taken for the first time, i.e. there are no previous observations.
  - *drop-off*: The scooter was left.
  - *drop-off outlier*: The scooter drop-off point is far from any placename ID.
  - *last seen / drop-off*: The scooter was left for the last time, i.e. there are no next observations.
  - *waiting long*: The scooter did not change position.

In summary, the dataset is a time series of pick-up and drop-off points for many different e-scooters.

4.1 Exploratory Data Analysis

First of all, we analyzed some basic statistical properties regarding pick-up and drop-off distances from the nearest marker (distance_p and distance_d) and time difference between observations, reported in table 1.

	Distance_p (m)	**Distance_d (m)**	**TimeDiff (min.)**
Mean	181.19	457.27	54.05
Median	180.34	13.59	15.02
Std. dev.	78.83	1029.11	275.98
Min	5.56	0.0	0.0
Max	762.84	7393.66	8864.92

Table 1: Basic statistics for distances and time

The median trip duration is 15 minutes, confirming the generally limited use of e-scooters. Nevertheless, the mean and standard deviation show the presence of several outliers, pushing the average value of each trip to 54 minutes. Also, looking at the values under the *distance_d* column, we notice the presence of several e-scooters left in far-flung areas of the city. These data confirm the largely recreational use of the vehicle, but also a fair amount of use for longer distances that would otherwise be covered by public transportation or private car. Finally, marker distances to pick-up points (*distance_p*) are less variable and with a much smaller maximum value than *distance_d*. E-scooters, upon being picked up, are occasionally abandoned in distant areas of the city, rendering them inaccessible to a significant number of users. Hence, a vehicle relocation plan becomes imperative. Similar observations can be drawn by examining figure 6, which provides a more detailed depiction of the variables distribution.

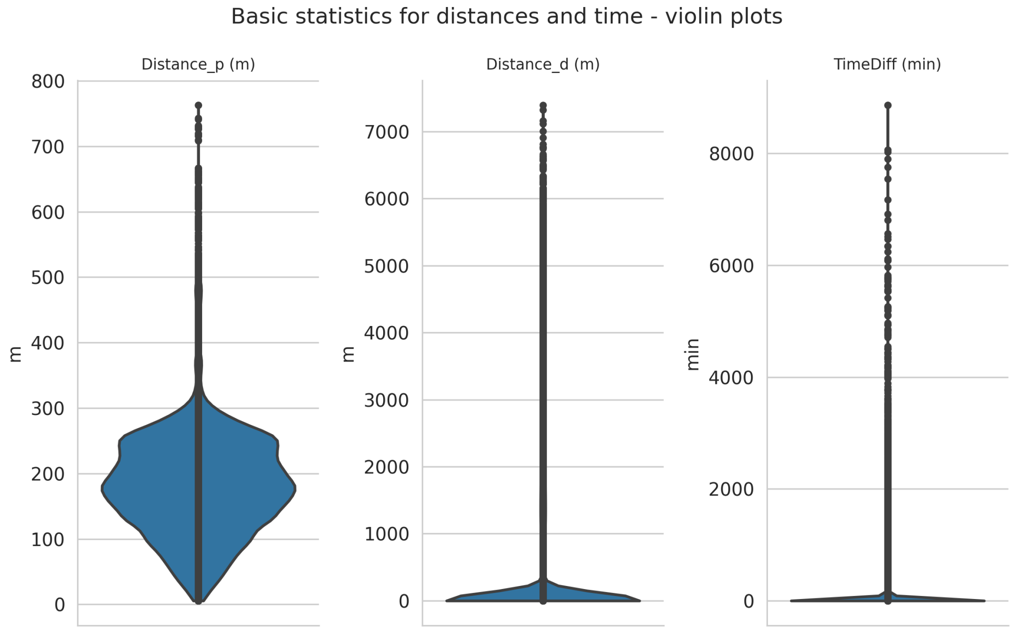


Figure 6: Basic statistics for distances and time - violin plots

During the conduct of the position prediction tasks, discretized coordinates were used (i.e. *placelat* and *placelon*), in combination with the Milan NIL subdivision. Figure 7 expands the previous choropleth map with the discretized location of e-scooters. From an overall view of the graph, e-scooters turn out to be in areas already well served by public transportation.

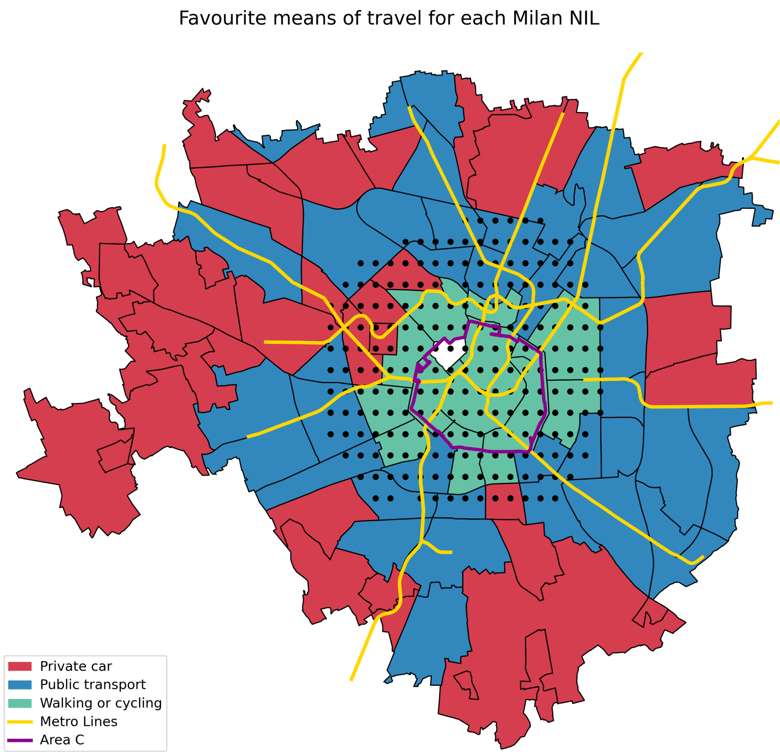


Figure 7: Means of transportation preferred by the citizens of Milan for each NIL in 2021 and e-scooters discretized positions in 2019

Finally, we created several choropleth maps showcasing the distribution of e-scooter locations. Figure 8 represents the total count of e-scooters observed throughout the week, while Figure 9 highlights the variation in usage between workdays and weekend.

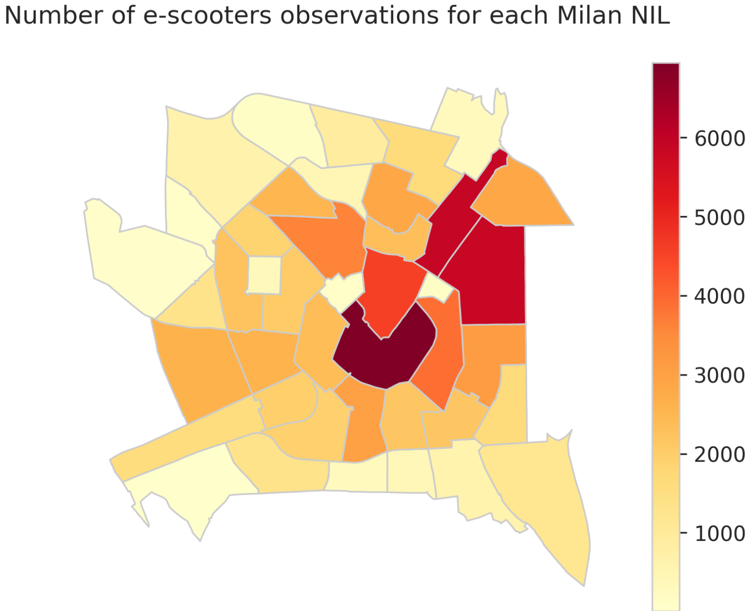
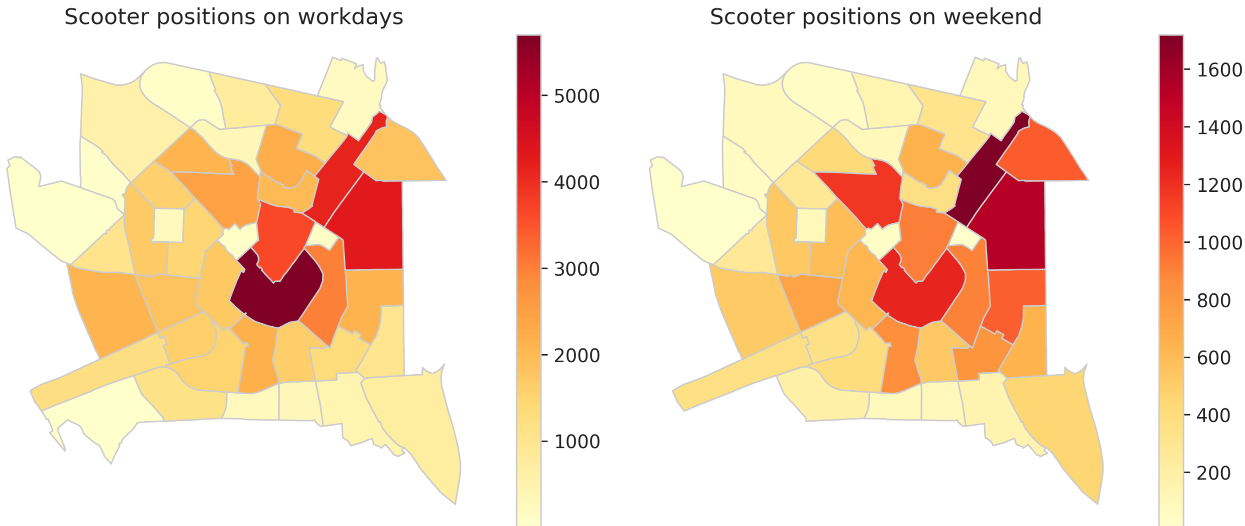


Figure 8: Number of e-scooters observations for each Milan NIL



(a) Workdays (Dec. 2-6)

(b) Weekend (Dec. 7-8)

Figure 9: Number of e-scooters observations for each Milan NIL

Undoubtedly, the busiest area in Milan is the cathedral (NIL *Duomo*), located at the heart of the city. However, there is a notable shift in patterns between workdays and the weekend. Specifically, on Saturday and Sunday, the nightlife districts like Porta Venezia, Sarpi and Porta Romana experience significantly higher activity. This surge can also be attributed to the subway closing around 00:30, prompting increased e-scooters usage in these areas. This also confirms the prevalent use by young people. Further details are provided by table 2.

General			Workdays			Weekend		
NIL	**Day count**	**%**	**NIL**	**Day count**	**%**	**NIL**	**Day count**	**%**
Duomo	992.29	8.40	**Duomo**	1139.80	9.17	**Centrale**	859.50	8.36
Centrale	837.71	7.09	**Porta Venezia**	616.29	6.94	**Porta Venezia**	761.50	7.41
Porta Venezia	833.86	7.06	**Centrale**	829	6.97	**Duomo**	623.50	6.07
Brera	651.71	5.52	**Brera**	728.80	5.86	**Sarpi**	586	5.70
Guastalla	556.29	4.71	**Guastalla**	598.20	4.81	**Loreto**	516.50	5.03
Sarpi	515.71	4.37	**Sarpi**	487.60	3.92	**XXII Marzo**	509.50	4.96
XXII Marzo	447.43	3.79	**Isola**	443.40	3.57	**Brera**	459	4.47
Ticinese	431.71	3.65	**Ticinese**	433	3.48	**Guastalla**	451.50	4.39
Isola	410.29	3.47	**XXII Marzo**	422.60	3.40	**Ticinese**	428.50	4.17
Loreto	409.86	3.47	**Ghisolfa**	418.20	3.37	**Porta Romana**	411	4.0
Tot	6086.86	51.53	**Tot**	6116.89	51.49	**Tot**	5606.5	54.56

Table 2: Top-10 busiest NIL: workdays and weekend

When examining the percentages in each area, it becomes evident that e-scooters traffic exhibits a greater polarization on workdays compared to the weekend. This can likely be attributed to the presence of diverse nightlife venues and the increased freedom workers have during the weekend. The data suggest that a significant portion of e-scooters users during the week employ the service for commuting to busy workplaces. Consequently, these users are less active on Saturday and Sunday, resulting in a reduced polarization on those days. In addition, e-scooters are generally more used during workdays. Therefore, for this particular case, e-scooters actually turn out to be a productive form of alternative mobility and not necessarily just an object of entertainment.

5 Problem definition

Considering the limited availability of specific indicators for this particular topic in the city of Milan, we decided to maximize the utilization of the existing data to create an index of "effective availability", which offers a comprehensive hourly overview of e-scooters availability in each area. The challenge of insufficient e-scooters presence in specific areas of the city during certain hours, caused by inefficient relocation practices, has already been acknowledged and addressed by researchers. For instance, the authors of [15] have proposed relocation strategies by employing deep learning techniques and simulations, while [13] and [14] have focused on demand forecasting. We opted for a different perspective, emphasizing the end user experience in tackling the issue.

To establish a foundation for addressing the problem, our initial approach involved estimating the hourly availability of e-scooters for each specific area. We organized the data by Neighborhood Identification Location (NIL) and hour. Starting from midnight, where an equal distribution of e-scooters across all zones has been assumed (i.e., 0 e-scooters for each zone), we tracked the number of e-scooters added (+1) or subtracted (-1) based on drop-offs or pick-ups within each zone, reflecting the changes in availability over time. However, this measure alone may not provide a comprehensive indication of e-scooter availability, as it does not consider the average demand in each specific area.

To address this limitation, we further estimated the average demand for each zone by calculating the average time interval between two consecutive pick-ups for each hour. By subtracting this value from the number of e-scooters remaining in the zone, we obtained a surplus/deficit rate. To create a standardized measure, we normalized this value between -1 and 1, resulting in an index that reflects the effective availability of e-scooters during the analyzed week.

In formula:

$$EA_{h,z} = RS_{h,z} - \frac{3600}{\mu_{h,z}(s)},$$

$$EAI_{h,z} = \frac{EA_{h,z} - min(EA_{h,z})}{2 \times (max(EA_{h,z}) - min(EA_{h,z})) - 1}, \text{ where:}$$

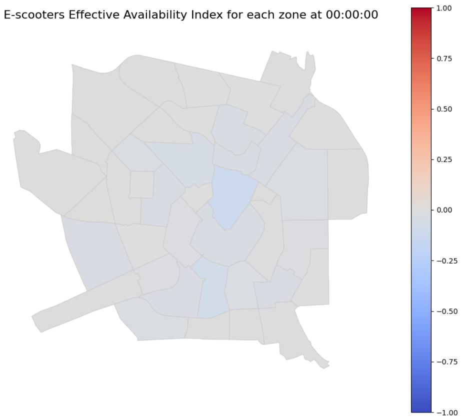
- $RS_{h,z}$ is the number of Remaining Scooters available in a specific hour h for a specific zone z.

- $\mu_{h,z}(s)$ is the average time (in seconds) between two consecutive pick-ups in a specific hour h for a specific zone z.

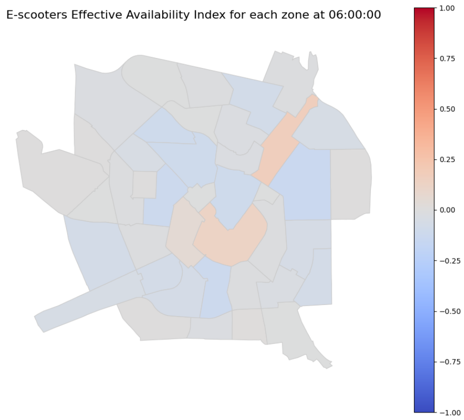
- $EA_{h,z}$ is the Effective availability, i.e. the number of e-scooters remaining minus the average number of e-scooters rented in a specific hour h for a specific zone z. This gives us the surplus or the shortage of e-scooters related to each specific hour and zone.

- $EAI_{h,z}$ is the Effective availability Index, i.e. the Effective Availability normalized between -1 and 1. The lower the index, the lower the Effective Availability of e-scooters.

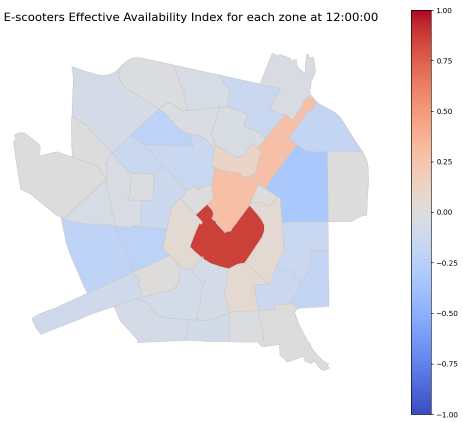
This index enabled us to analyze the actual surplus or deficit of e-scooters in each zone for every hour (fig. 10). An animated version of this and other choropleth maps is available by clicking on [24].



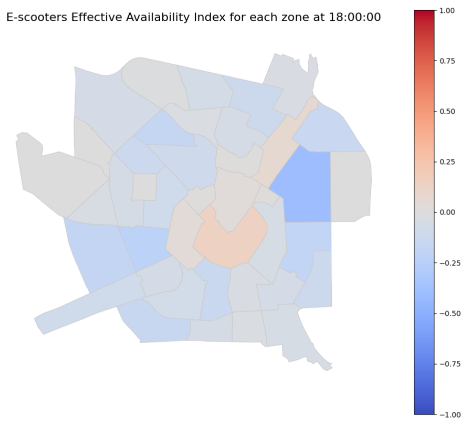
(a) H0



(b) H6



(c) H12



(d) H18

Figure 10: Effective availability Index

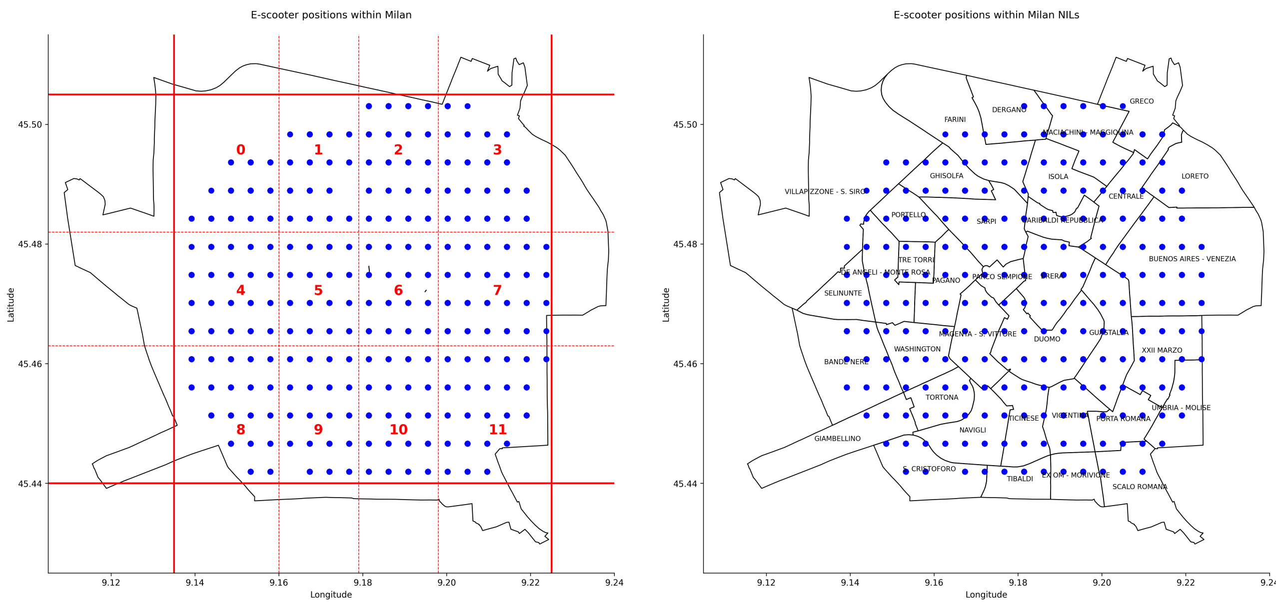
Throughout the day, e-scooters availability fluctuates, resulting in varying surpluses and deficits. At 12 noon, an interesting pattern emerges with the Duomo area exhibiting a significant surplus, while several other areas experience deficits. Notably, Buenos Aires - Porta Venezia has the highest deficit (EAI = -0.32), a situation that persists at 6 p.m. (EAI = -0.41). Additionally, Ghisolfa, Washington, and Bande Nere also display indices below -0.2 at noon. Although the surplus decreases by 6 p.m., the EAI for these areas only slightly improves, indicating an ongoing shortage of e-scooters.

In the following section, we present an alternative solution that aims to assist end users in their decision-making process when selecting an e-scooter. Moreover, by predicting the potential drop-off zones, both providers and public authorities can gain a deeper understanding of the phenomenon and develop effective relocation strategies. The forecasted information would enable them to identify areas with the highest supply in advance, providing them with more time to allocate additional resources and ensure an adequate availability of e-scooters in areas with a service deficit. This proactive approach might contribute to a more balanced distribution of e-scooters and offers a practical approach to meet users needs.

6 Proposed LSTM model

6.1 Research question #1

To thoroughly explore the potential of this approach, a series of graded assessments has been conducted, encompassing a range of task difficulties and rationales. We therefore started with the first research question: "Can the same drop-off zone for a sequence of e-scooters be predicted by analyzing their individual trajectories?", i.e. given a certain number of e-scooters, each with a certain trajectory but the same drop-off point, is it possible to predict their collective arrival zone? The task, a multi-class classification, was performed by varying the number of zones to be predicted and the length of the input e-scooters sequence. Regarding zoning, a join was made between the discretized location of the e-scooters (*placelat* and *placelon* columns) and the subdivision of the city of Milan. Specifically, a geographic partition into 12 discrete zones was first made (fig. 11a), then the task was complicated by directly using the division into NILs (fig. 11b). As for the input sequence, the length tested was 10 and 5 e-scooters.



(a) Partition into zones - 12 classes

(b) Partition into NILs - 36 classes

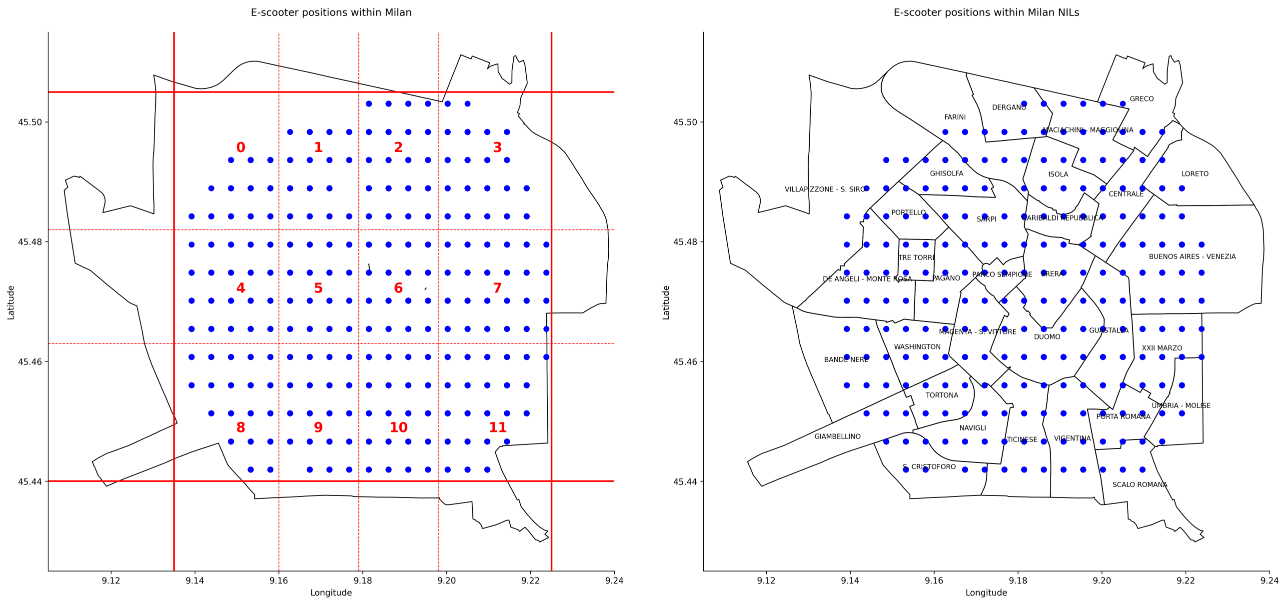
Figure 11: Geographical partitions for research question #1

Naturally, in both cases, the research boundaries were defined based on the presence of e-scooters. Furthermore, for the division into NILs, it is important to clarify that due to limited data availability in *QT 8*, *Giardini di Porta Venezia*, *Bovisa*, *Città Studi* and *San Siro*, these areas were merged with neighboring regions, respectively *Villapizzone*, *Guastalla*, *Farini*, *Buenos Aires - Porta Venezia* and *Villapizzone* again, renamed to *Villapizzone - San Siro*. That being said, the total number of classes considered in this context was 36.

In summary, for this research question, we generated a total of four different models by combining the two sequence lengths (10 and 5) and the two class divisions (12 and 36).

6.2 Research question #2

We then moved on to the second research question: "Can the drop-off zone of each individual e-scooter be predicted by analyzing its individual trajectory?", i.e. given a certain trajectory of the same e-scooter, is it possible to predict its arrival point? In contrast to the previous approach, the dataset was additionally grouped by the ID of individual e-scooters, resulting in a higher level of prediction granularity. In this case, we no longer considered groups of e-scooters with the same drop-off point. Instead, each e-scooter has been treated individually, considering its unique trajectory. Due to this additional grouping, three zones, namely *Ex Om - Morivione*, *Selinunte* and *Tibaldi*, did not have sufficient data to be considered as separate entities for classification on NILs. As a result, these zones were merged with *Vigentina*, *De Angeli - Monte Rosa*, and *Ticinese*, respectively. The total number of classes considered thus became equal to 33. For this task, we opted to consider a sequence length of 5 e-scooters. This decision was made to mitigate the loss of data, as a longer sequence length would have resulted in a greater data loss. In addition, it is worth noting that shorter sequence lengths tend to make the prediction task more complex for the model. If the network demonstrates good performance with a sequence length of 5, there is potential for further improvement by increasing the length of the sequence. Figure 12 shows the partitions for this second multi-class classification task.



(a) Partition into zones - 12 classes

(b) Partition into NILs - 33 classes

Figure 12: Geographical partitions for research question #2

In summary, for this research question, we generated two different models by combining the two class divisions (12 and 33) with a sequence length of 5 e-scooters.

In the following subsections, we outline the technical methodology employed, which remains consistent for both tasks, with a few exceptions.

6.3 Data preparation

To derive the trajectories of each e-scooter, the dataset was initially grouped by e-scooter ID and, within each group, the observations were sorted chronologically based on time. The dataset was subsequently divided into pick-up observations and drop-off observations. Observations where the type column indicated "waiting long" were excluded from the analysis due to their limited occurrence and little significance for trajectory purposes. A new dataset was then constructed to consolidate the drop-off information into the corresponding pick-up rows. After selecting the relevant columns, the resulting dataset took the following form:

- *provider_id*: e-scooter ID.
- *created_at*: e-scooter pick-up timestamp.
- *placelon*: e-scooter pick-up longitude.
- *placelat*: e-scooter pick-up latitude.

- *is_weekend*: binary variable indicating whether the observation was taken over the weekend, when there is a slight change in trajectories compared to workdays.

- *target*: the drop-off point (zone or NIL).

For the purposes of the first research question, the dataset was then grouped by target and each group was reordered temporally, regardless of the ID of the e-scooters, while for the second research question this aspect was also considered. Figure 13 shows the difference between the two datasets across the first 5 observations, considering the NIL classification.

	provider_id	created_at	placelon	placelat	is_weekend	target
0	1342.0	2019-12-02 00:10:11	9.13915	45.46545	0	BANDE NERE
1	1285.0	2019-12-02 00:25:11	9.13915	45.46545	0	BANDE NERE
2	1342.0	2019-12-02 00:40:09	9.13915	45.46545	0	BANDE NERE
3	1373.0	2019-12-02 00:40:09	9.14385	45.46545	0	BANDE NERE
4	1098.0	2019-12-02 00:40:09	9.14385	45.46545	0	BANDE NERE

(a) Dataset for research question #1

	provider_id	created_at	placelon	placelat	is_weekend	target
0	708.0	2019-12-03 08:55:10	9.14855	45.45605	0	BANDE NERE
1	708.0	2019-12-03 09:25:10	9.14855	45.45605	0	BANDE NERE
2	708.0	2019-12-03 11:10:10	9.14855	45.45605	0	BANDE NERE
3	708.0	2019-12-04 00:25:10	9.14855	45.45605	0	BANDE NERE
4	708.0	2019-12-04 22:10:08	9.14855	45.45605	0	BANDE NERE

(b) Dataset for research question #2

Figure 13: First 5 observations from the two datasets for NIL classification

The dataset was subsequently partitioned into training, validation and test. Considering the lower reliability of an anomalous day such as Sunday, and in light of the additional factor of the feast of the Immaculate Conception (December 8), the first day of data, which was Monday, December 2, was selected as the test set. This choice was made under the assumption that this day would exhibit a similar trend to that of the following week (Monday, December 9). In contrast, for the validation set, a different approach was adopted. In the case of the first dataset, the last 10% of each target was designated, while for the second dataset it was necessary to select the last 30% of each trajectory to ensure proper representation for all the target classes in the validation set. We show an illustrative example of the dataset partitioning in figure 14 for ease of understanding.

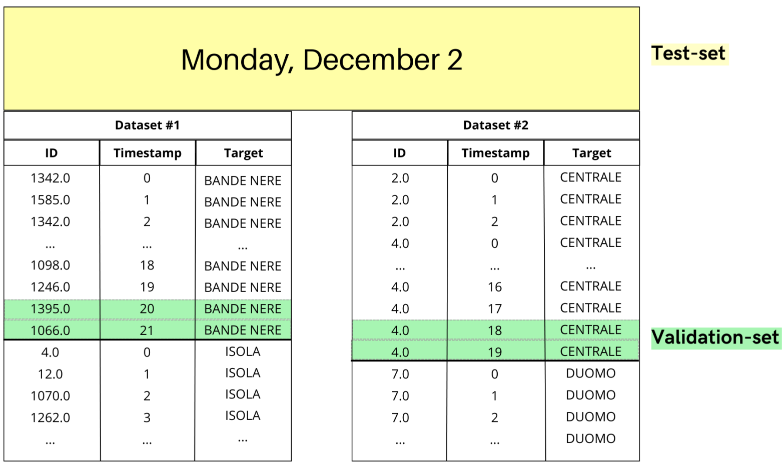


Figure 14: Example of split into training, validation and test

To incorporate temporal information into the model, the *created_at* column was converted to Unix time. Subsequently, each element in the column was subtracted by the first Unix time value in the training set. This transformation ensured that the temporal count started from 0, enabling the model to capture the relative time differences between data points effectively.

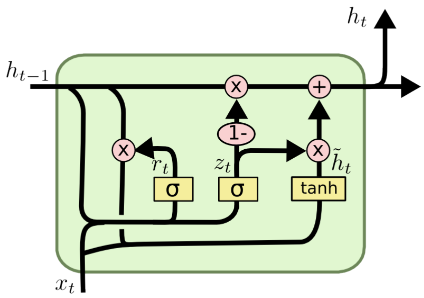
As for *provider_id* and *target*, these were encoded using the *LabelEncoder* function from the *scikit-learn* library [25]. This encoding process assigns unique integer values to each distinct category, facilitating the numerical representation of categorical variables in the dataset.

In order to ensure uniformity and address stability issues arising from variables with different scales, each input variable was scaled to a range between 0 and 1. This process maintains the unique encoding of categorical variables, as decimal values between 0 and 1. Moreover, it helps mitigate any instability that may arise from using weights with varying scales, promoting a more balanced and reliable model performance.

Lastly, the data were divided into sequences of either 10 or 5 observations, based on the specific task and research question at hand.

6.4 Model architecture

Recurrent Neural Networks (RNNs), particularly those equipped with Long Short-Term Memory (LSTM) units [26], are widely employed for time series forecasting tasks. These models are preferred because they excel at capturing long-term dependencies in sequential data. LSTMs are a specialized variant of RNNs that address the limitations of traditional RNNs, as they possess mechanisms to mitigate the issues of vanishing or exploding gradients. By utilizing memory cells and gating mechanisms (see fig. 15), LSTMs can selectively retain or forget information over extended sequences, allowing them to capture and retain relevant contextual information over time. The inclusion of LSTM units in our neural network architecture enhances its capability to capture and leverage long-term dependencies in the time series data. This is particularly advantageous for accurate and reliable forecasting, as the model can effectively leverage historical information to make more informed predictions.



$$z_t = \sigma\left(W_z \cdot [h_{t-1}, x_t]\right)$$
$$r_t = \sigma\left(W_r \cdot [h_{t-1}, x_t]\right)$$
$$\tilde{h}_t = \tanh\left(W \cdot [r_t * h_{t-1}, x_t]\right)$$
$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t$$

Figure 15: LSTM unit architecture

Nevertheless, we decided to maintain a consistent and simple structure across all tasks, in order to enhance the efficiency and robustness of the model. The overall architecture of the model consists of two stacked LSTM layers, which are followed by two fully-connected layers (see fig. 16). The final fully-connected layer outputs a probability vector using the softmax activation function [27]. By adopting a uniform architecture, we ensure a streamlined and standardized approach throughout the various tasks. This approach simplifies the model development and training process, as well as facilitates comparisons and evaluations across different tasks.

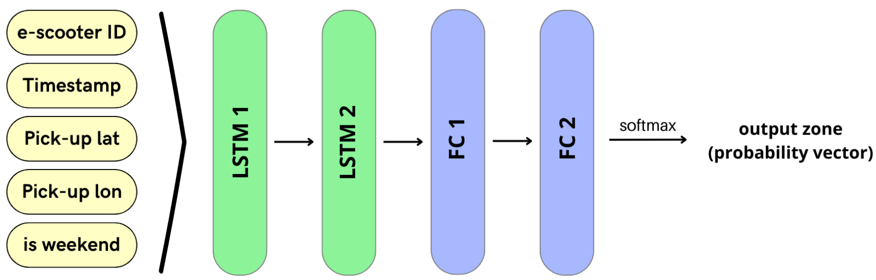


Figure 16: General model architecture

6.5 Model training & tuning

The models were built, trained, and tuned using the popular *Keras* [28] library, a high-level deep learning API implemented in Python and designed to run on top of the *TensorFlow* [29] machine learning platform.

Specifically, we conducted an iterative model selection process for each task, where we experimented with different hyperparameters, using the *GridSearch* functionality provided by *KerasTuner* [30]. The goal was to identify the optimal combination that yielded the best performance on the validation set. Once the best performing models (one for each task) were determined, they have been re-initialized and re-trained using the entire dataset, excluding the test set, and inferring the number of epochs from the loss trend on the validation set. By re-training each model on the complete dataset, we aimed to leverage as much data as possible to improve the performances and generalization capabilities of the models. Finally, after the re-training phase, each model has been evaluated through the reserved test set. This independent evaluation allowed us to assess the model accuracy on unseen data.

Table 3 provides a summary of the different combinations of hyperparameters that were explored during the model selection processes. The best hyperparameter configuration for each model is reported in the pertaining column. It is important to note that all models were trained for 200 epochs with early stopping and a patience of 30 epochs, which would stop the training if there was no improvement in validation loss for 30 consecutive epochs, preventing overfitting. Additionally, the "restore best weights" approach was employed, which ensures that the weights corresponding to the lowest validation loss are restored before evaluation. Regarding the parameters of the model compile function, we employed the Adaptive Moment Estimation (*Adam*) optimizer and the categorical crossentropy as loss function.

	Research question #1				**Research question #2**	
Hyperparameters	Classes: 12 Length: 10	Classes: 12 Length: 5	Classes: 36 Length: 10	Classes: 36 Length: 5	Classes: 12 Length: 5	Classes: 33 Length: 5
LSTM1 units [16, 32]	32	32	16	32	32	32
LSTM2 units [32, 64]	32	32	32	32	64	32
FC1 units [32, 64]	32	32	64	32	32	32
Learning Rate [0.0001, 0.001]	0.001	0.001	0.001	0.001	0.001	0.001
Dropout LSTM1 [0, 0.2]	0.2	0.2	0.2	0.2	0.2	0.2
Dropout LSTM2 [0, 0.2]	0.2	0.2	0.2	0.2	0.2	0.2
Dropout FC1 [0, 0.2, 0.4]	0.2	0.4	0.4	0.2	0.2	0.2
KernelReg LSTM1 [0, 0.01]	0	0	0	0	0	0
BiasReg LSTM1 [0, 0.01]	0.01	0	0.01	0	0.01	0.01
KernelReg LSTM2 [0, 0.01]	0	0	0	0	0	0
BiasReg LSTM2 [0, 0.01]	0	0	0	0	0	0
KernelReg FC1 [0, 0.01]	0	0	0	0	0	0
BiasReg FC1 [0, 0.01]	0.01	0.01	0	0	0	0
Batch size [16, 32]	16	32	16	32	32	32
Val accuracy	97.7%	93.9%	90.8%	84.3%	99.2%	97.9%

Table 3: Hyperparameters fine-tuning - KerasTuner GridSearch

Hyperparameters tuning resulted in significant improvements for all tasks. In particular, the addition of dropout layers proved particularly effective. Dropout [31] is a regularization technique commonly used in neural networks, which helps to prevent overfitting and improves the generalization ability of the model. In dropout, during training, a certain fraction of the neurons in a layer (i.e. the parameter of the function) is randomly selected and set to zero with a given probability. This means that the dropped out neurons do not contribute to the forward pass of the network and are not updated during backpropagation. By randomly dropping out neurons, this technique forces the remaining neurons to learn more robust and independent representations of the data. This reduces the reliance of the network on specific neurons and mitigates overfitting. The introduction of L2 regularization, which adds a square penalty term to the loss function, has also been shown to improve performance by further reducing overfitting. Lastly, it is interesting to observe that longer input sequences prefer smaller batch sizes.

7 Results & discussion

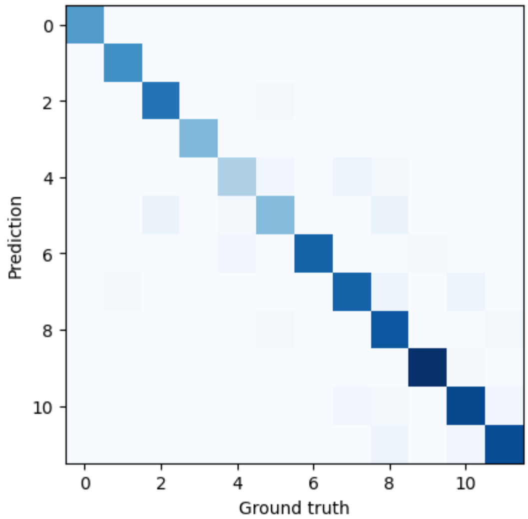
Finally, the generalization capabilities were evaluated for each of the six models by comparing, on the test set, the predicted classes with the ground truth. Table 4 summarizes the results achieved in terms of accuracy, precision and recall, while figure 17 shows the distribution of predictions and ground truth across classes, for all tasks.

	Research question #1				**Research question #2**	
Metrics	Classes: 12 Length: 10	Classes: 12 Length: 5	Classes: 36 Length: 10	Classes: 36 Length: 5	Classes: 12 Length: 5	Classes: 33 Length: 5
Top-1 accuracy	93%	87%	84%	78%	99%	93%
Top-2 accuracy	98%	94%	92%	87%	99%	99%
Precision	93%	87%	84%	77%	98%	92%
Recall	93%	87%	83%	77%	98%	92%

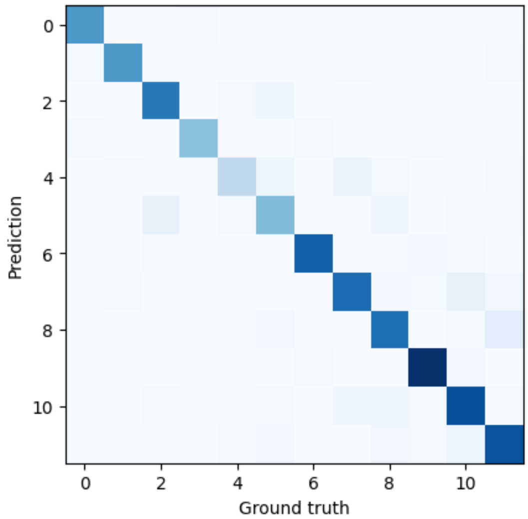
Table 4: Test set evaluation: accuracy, precision, recall

The efficiency of the models is generally satisfactory for all tasks. For the first research question, performances naturally degrade as task complexity increases. However, even in the most difficult case, the model succeeds in predicting the exact zone 78% of the time and, in 87% of the cases, the NIL is correctly predicted by the first two guesses. On the other hand, for the second research question, the accuracy, precision and recall values are surprisingly high for both tasks. Thus, grouping with respect to the e-scooter ID turned out to be a key piece of information. Receiving sequences of single trajectories as input, the model achieves more than 90% prediction accuracy, even with the shortest sequence length of 5.

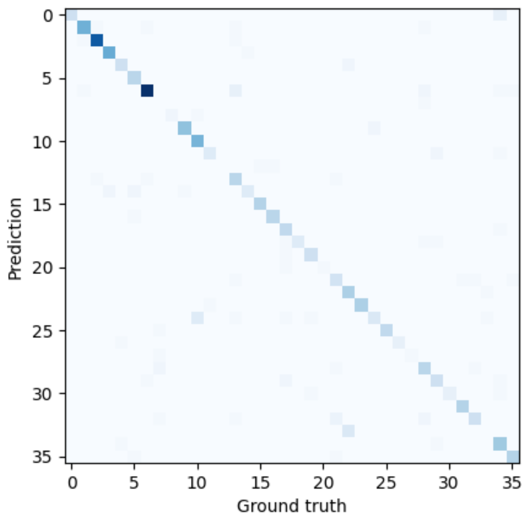
In summary, despite only one week's worth of data, we obtained results that would allow practical application for both research questions. By having more days of data, the models could be further improved and, in addition, they would be able to capture weekly seasonality. Specifically, the accurate prediction of the drop-off zone for individual e-scooters, as explored in research question #2, has the potential to greatly enhance the user mobility experience if integrated as an additional feature in existing apps offered by various service providers.



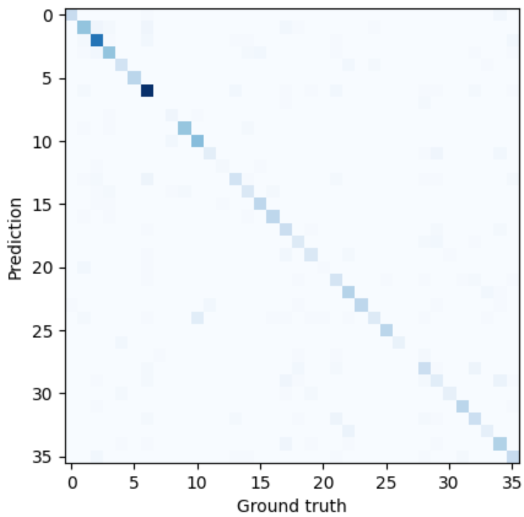
(a) RQ #1: classes: 12, length: 10



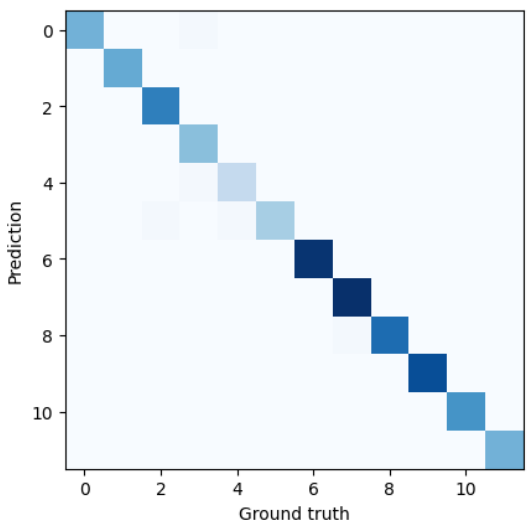
(b) RQ #1: classes: 12, length: 5



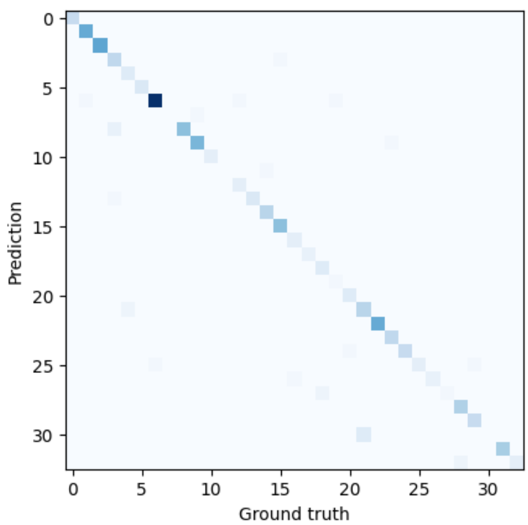
(c) RQ #1: classes: 36, length: 10



(d) RQ #1: classes: 36, length: 5



(e) RQ #2: classes: 12, length: 5



(f) RQ #2: classes: 33, length: 5

Figure 17: Predictions vs. ground truth

8 Conclusions

In this research, we demonstrated the effectiveness of a light neural network with LSTM units for predicting the drop-off zones of e-scooters in the city of Milan.

To begin, we framed our study within the broader context of smart cities and smart mobility, focusing specifically on the city of Milan. After analyzing the available dataset, we approached the problem by creating a customized Effective Availability Index (EAI). This index allowed us to identify areas in the city that exhibited either a surplus or deficit of e-scooters throughout the hours of the day.

In order to address the two research questions, we employed a consistent methodology while adapting the data grouping. This involved training six models with similar architectures, each tailored to the length of the e-scooters sequences to be analyzed and the desired number of classes (i.e. the different zones of Mlan). Grid search was finally employed to optimize the hyperparameters for each specific task.

The obtained results consistently showcase the effectiveness of our approach across all conducted tasks. The accuracy of the predictions generally exceeds 80%, with peaks over 90% and precision and recall values in a similar range. These findings hold promising practical implications, as they could greatly enhance the management of e-scooters fleet by both public authorities and service providers. Simultaneously, they offer the potential to significantly improve the overall user experience.

8.1 Future work

We believe that despite its demonstrated effectiveness, there are several avenues for further expansions and improvements in this research. While we deliberately chose a lightweight approach during the design stage, there is potential for exploring more complex models, although this would need to be carefully balanced.

One significant improvement could be achieved by increasing the amount of available data. By incorporating a month's worth of data, for instance, the model could capture additional patterns and weekly seasonality, thereby enhancing the accuracy of zone predictions. Furthermore, with a larger dataset, it would be possible to analyze sequences of 10 e-scooters for research question #2, offering a more comprehensive perspective.

Additionally, conducting the research using more recent data (e-scooters observations date back to December 2019) or data from other categories of sharing mobility (e.g. cars, bikes) might provide valuable insights and a broader understanding of the phenomenon.

Lastly, the robustness of the approach may be tested by introducing the estimation of drop-off time alongside the area prediction, thereby adding an additional level of complexity to the task.

Overall, we think that these suggested improvements have the potential to further enhance the research outcomes and contribute to a more comprehensive understanding of e-scooters dynamics and mobility patterns.

References

[1] L. Sartori and D. Arcidiacono, "In search for (the lost) smartness in smart cities: Consumers or citizens?" 01 2021.

[2] D. Grimaldi and C. Carrasco-Farré, *Implementing Data-Driven Strategies in Smart Cities: A Roadmap for Urban Transformation*, ser. Smart Cities Series. Elsevier Science, 2021. [Online]. Available: https://books.google.it/books?id=_xkJzgEACAAJ

[3] (2015) European smart cities 4.0. [Online]. Available: https://www.smart-cities.eu/?cid=01&ver=4

[4] M. Foth, J. Choi, and C. Satchell, "Urban informatics," 03 2011, pp. 1–8.

[5] R. Fistola, M. Gallo, and R. A. La Rocca, "Cities between smartness and emergencies: exploring the role of e-scooter in the "transition era"," 01 2021.

[6] M. Fazio, N. Giuffrida, M. Le Pira, G. Inturri, and M. Ignaccolo, "Planning suitable transport networks for e-scooters to foster micromobility spreading," vol. 13, no. 20, 2021.

[7] (2021) Milano smart city alliance. [Online]. Available: https://milanosmartcity.it/istitutional/

[8] D. Tamiozzo and M. Colleoni, "E-scooter accidents and the urban environment: the case of milan," 06 2021. [Online]. Available: https://transformtransport.org/research/innovative-technologies/e-scooter-accidents-and-the-urban-environment-the-case-of-milan/

[9] N. Cornali, M. Seminati, P. Chiodini, and P. Maranzano, "Smart mobility in milan, italy: A district-level spatial and cluster analysis," vol. 34, 02 2023.

[10] V. Negri, F. Coppola, P. Testa, M. Mena, F. Zoppis, and C. Quintiliano, "Booklet smart city: Milano nel confronto europeo," 2020.

[11] A. Li, P. Zhao, X. Liu, A. Mansourian, K. W. Axhausen, and X. Qu, "Comprehensive comparison of e-scooter sharing mobility: Evidence from 30 european cities," vol. 105, 2022, p. 103229.

[12] A. Sellaouti, O. Arslan, and S. Hoffmann, "Analysis of the use or non-use of e-scooters, their integration in the city of munich (germany) and their potential as an additional mobility system," in *2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)*, 2020, pp. 1–5.

[13] A. Ciociola, M. Cocca, D. Giordano, L. Vassio, and M. Mellia, "E-scooter sharing: Leveraging open data for system design," 09 2020, pp. 1–8.

[14] S. W. Ham, J.-H. Cho, S. Park, and D.-K. Kim, "Spatiotemporal demand prediction model for e-scooter sharing services with latent feature and deep learning," vol. 2675, 04 2021.

[15] L. Tolomei, S. Fiorini, A. Ciociola, L. Vassio, D. Giordano, and M. Mellia, "Benefits of relocation on e-scooter sharing - a data-informed approach," in *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*, 2021, pp. 3170–3175.

[16] J. Violos, S. Tsanakas, M. Androutsopoulou, G. Palaiokrassas, and T. Varvarigou, "Next position prediction using lstm neural networks," 09 2020, pp. 232–240.

[17] D. Bongardt, L. Stiller, A. Swart, and A. Wagner, "Sustainable urban transport: Avoid-shift-improve (a-s-i)," 2019.

[18] C. Vitrano and M. Colleoni, *Mobility as a Service: Moving in the De-synchronized City*, 06 2020.

[19] (2022) Analisi della variazione degli indicatori relativi ai sistemi di mobilità di milano. [Online]. Available: https://datashare.amat-mi.it/index.php/s/HfH8so2oSCMEbB6?dir=undefined&path=%2F2022&openfile=332317

[20] M. Ciuffini, S. Asperti, V. Gentili, R. Orsini, and L. Refrigeri, "6° rapporto nazionale sulla sharing mobility," 2022.

[21] (2019) Parco veicolare circolante a milano. [Online]. Available: https://dati.comune.milano.it/it/dataset/ds1093-parco-veicolare-dei-residenti-a-milano-suddiviso-per-categoria-anno-2019/resource/575ce1b0-b4e6-4222-bd79-178422af228f

[22] (2021) Mezzo di trasporto prevalente a milano. [Online]. Available: https://dati.comune.milano.it/dataset/ds1490-mezzi-di-trasporto-di-utilizzo-prevalente/resource/ab30907e-47b8-45e2-a305-00fe6d4b3e90

[23] (2021) Parco sharing a milano - serie storica. [Online]. Available: https://dati.comune.milano.it/dataset/ds574-servizi-di-car-sharing-e-bike-sharing/resource/e3a3d46e-9b52-41ac-ae48-789b44f9c58e

[24] Animated choropleth maps. [Online]. Available: https://drive.google.com/drive/folders/15GOu0pCDG2ZTrobwwQDahCJHHGPHEA4Q?usp=share_link

[25] Scikit-learn library. [Online]. Available: https://scikit-learn.org/stable/

[26] S. Hochreiter and J. Schmidhuber, "Long short-term memory," vol. 9, 12 1997, pp. 1735–80.

[27] J. Bridle, "Training stochastic model recognition algorithms as networks can lead to maximum mutual information estimation of parameters," in *Advances in Neural Information Processing Systems*, D. Touretzky, Ed., vol. 2. Morgan-Kaufmann, 1989.

[28] Keras library. [Online]. Available: https://keras.io/

[29] Tensorflow platform. [Online]. Available: https://www.tensorflow.org/?hl=it

[30] Keras tuner. [Online]. Available: https://keras.io/keras_tuner/

[31] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," vol. 15, no. 56, 2014, pp. 1929–1958.