Parking Demand Forecast using Additive Model: a proof of concept on a real parking meters dataset

Luca Romeo  
Department of Information Engineering  
Università Politecnica delle Marche  
Ancona, Italy  
l.romeo@univpm.it

Davide Alessandrini  
Pluservice srl  
Senigallia, Italy  
d.alessandrini@pluservice.net

Eleonora Bicchierini  
Pluservice srl  
Senigallia, Italy  
eleonora.bicchierini@mycicero.it

Emanuele Frontoni  
Department of Information Engineering  
Università Politecnica delle Marche  
Ancona, Italy  
e.frontoni@univpm.it

Abstract

The recent advance in AI and the increasing amount of available data move us closer to developing an urban operating system that provides sustainable and inclusive mobility, a multimodal public transport system and accurate monitoring of traffic and parking congestion. Thus, accurate and timely information about parking occupancy and availability has played a crucial role to solve the smart city challenge related to mobility, by helping drivers to save their time and by avoiding waiting to find an empty space, to move smoothly, or be in traffic. In particular, the design and the development of a parking recommendation system (PRS) to create real-time parking and traffic maps may answer the increasing demand to get available and timely parking space. At the same time, PRS based on machine learning and deep learning algorithm may improve parking forecasting by generalizing across different environments, weather, traffic and temporal conditions. In this work, we extend the concept of PRS by proposing the application of an additive regression model (Prophet) fed by parking meters data (parking meters amount). We tested the proposed model on a real parking meter dataset collected from Bologna city. The obtained forecasting results (overall mean absolute error up to 0.61€) suggest that the proposed approach is a viable solution for providing reliable forecasting of parking occupancy for different are by modeling non-linear, non-periodic and weekly periodic changes of the parking meter data. Moreover, the ML model is consistent with the numerical challenges which are typical of this data including missing data, outliers and dramatic variation in the time-series data.

Introduction

A smart city connects human capital, social capital and Information and Communication Technologies (ICT) infrastructure in order to address public issues, achieve sustainable development and increase the quality of life of its citizens. The smart city goals include but are not limited to (i) achieve sustainable development, (ii) increase the quality of life of its citizens and (iii) improve the efficiency of the existing and new infrastructures. The main actors can be defined as government and city authorities, public-private partnerships and citizen participation. Thus, the scope of the smart city is focused on the economic and social expansion of the whole region, the connected settlements and the improvement to the regional and interregional networks.

The recent advance in AI and the increasing amount of available data move us closer to developing an urban operating system that simulates human, machine, and environmental patterns. (Lim, Kim, and Maglio 2018). Smart Cities and AI brings together a multidisciplinary, integrated approach related to how the combination of human and machine intelligence is transforming the experience of the urban environment. The AI algorithm may provide an affordable solution to support cities to become self-regulated by exploring cities as real-time, living, dynamics system. One of the main challenges in this context is to improve mobility by providing sustainable and inclusive mobility, a multimodal public transport system and accurate monitoring of traffic and parking congestion.

Nowadays, the design and the development of a parking recommendation system (PRS) to create real-time parking and traffic maps may answer the increasing demand to get available and timely parking space. Thus, accurate and timely information about parking occupancy and availability has played a crucial role to solve the smart city challenge related to mobility, by helping drivers to save their time and by avoiding waiting to find an empty space, to move smoothly, or be in traffic. Accordingly, in the mobility scenario, a huge amount of data is now collected, not only from commercial organizations and public administrations but also from billions of individual ICT users.

In recent times, there has been growing interest in the use of Big Data and crowdsourcing data for both research and commercial applications. Research is progressing at great steps in the extraction of useful geospatial information from large social media datasets to support disaster management and the updating of open data such as OpenStreetMap.[1] However, several challenges still remain in order to extract salient information for designing an accurate and timely PRS. Differently from PRS mainly based on knowledge-based and rule-based approaches, the introduction of Ma-
Machine Learning and Deep Learning may facilitate the extraction of discriminative information for the detection of parking occupancy and behavior. These predictive models should be able to generalize across the different environments, weather, traffic and temporal conditions by providing a tempestively forecast of parking occupancy on large areas (e.g. city).

In the literature, the implementation of PRS based on ML has been of increasing interest possibly a result of the increasing availability of datasets. In (Wu et al. 2014), the authors developed an optimal recommendation sequence considering successful parking probability and time to reach the parking lot. As a case study, they considered the collection of realistic records from parking lots in Taipei city. Their predictive algorithm was designed to estimate the successful parking probability by using currently available space counts. Similarly, in (Vlahogianni et al. 2016), a PRS was presented in order to provide the probability of free space to continue being free in subsequent time intervals, and the short-term parking occupancy forecasting in selected regions of an urban road network. A standard regression model was used and the input data was collected by a wide network of on-street parking sensors in the “smart” city of Santander, Spain. In (Richter, Di Martino, and Mattfeld 2014), an unsupervised clustering algorithm was proposed to model the spatial and temporal representation of parking space availability by evaluating the parking availability data, publicly available from the project Spark, based in San Francisco. The PRS proposed in (Saharan, Kumar, and Bawa 2020) used a ML approach to forecast occupancy of parking lots, which in turn is used to forecast occupancy driven prices for arriving vehicles. They used the historical on-street parking data of Seattle City.

Differently from the current state of the art work our PRS extend the application of standard ML approaches by proposing the application of an additive regression model (Prophet) fed by parking meters data (parking meters amount). The proposed PRS provides timely forecasting of the next 24-h parking occupancy in terms of parking meters amount. The main advantages of the proposed additive model with respect to other state of the art approaches (e.g. (Yang et al. 2019)) which employ sequential deep learning model (e.g. LSTM) are the greatest flexibility (i.e. we can easily accommodate parking occupancy seasonality with multiple periods and let the analyst make different assumptions about trends), the robustness against missing values (i.e. no need to interpolate missing values), the low computation effort and the higher model interpretability. Moreover, the task definition as well as the implementation of the proposed PRS as the basis of the overall project originated from a specific company’s demands (Pluservice srl). As a result, the main contributions of this work to the field of Urban Mobility and AI are summarized below:

- the collection of a real parking meters dataset from Bologna city;
- the application of an additive regression model (Prophet) for forecasting the next 24-h parking occupancy;
- the experimental validation and testing of the proposed approach in different city areas.

**Dataset**

The dataset *Parking meters amount* employed in this study consists of all parking meters amount for each different area of Bologna city (see Figure 1) during 2019. All the dataset comprised of 4.387 mln observations.

![Figure 1: Parking meter positions in the city of Bologna](image)

Data are anonymous in terms of customers and instrument and their use, detention and conservation are regulated by an agreement between Pluservice company, Università Politecnica delle Marche and data owners.

**Task Definition**

Our objective is to forecast the parking occupancy in terms of parking meter amount for each different area. Thus, we aim to forecast the amount at time $t + 1$ starting from the amount at time $t$ (see Figure 2). We solve this task independently for each parking meter area, by considering the next day as the horizon of forecast ($t = i - th$ day).

![Figure 2: Task definition: parking meter amount forecasting](image)

**Method**

We propose the application of an additive model (Prophet) for forecasting the next 24-h parking occupancy in terms of parking meters amount (Gaillard, Goude, and Nedellec 2016; Taylor and Letham 2018). The rationale motivation
behind the application of this ML model is mainly due to the high robustness of the model to provide accurate forecasting by being robust to outliers, missing data and dramatic changes in the time series that represent the main three challenges of this task. The employed additive model is conceived for forecasting time series data where non-linear trends are fit with yearly, weekly, and daily seasonality, plus holiday effects. The Prophet is similar to a generalized additive model (GAM), with time as a regressor and is able to fit several linear and non-linear functions of time as components. The additive approach tries to model the following three components:

\[
\text{amount}(t) = g(t) + s(t) + e(t) \tag{1}
\]

where \(g(t)\) models non-periodic changes, \(s(t)\) models seasonality periodic changes (i.e. weekly, monthly) and \(e(t)\) models the idiosyncratic error. Notice how the Prophet model is essentially conceived to solve the forecasting problem as a curve-fitting task without looking explicitly at the time-based dependence of each observation. In our task, the \(g(t)\) is set to model a non-linear saturating trend using a logistic saturating growth model, while the \(s(t)\) is set to model the weekly effect with a standard Fourier series with a period equal to 7.

**Results**

**Pre-processing**

Afterward, the number of occurrences for each different area were evaluated. Figure 3 shows the top 20 ranking areas in terms of the number of occurrences (i.e. number of parking meter occurrences) during all year.

![Figure 3: Number of observations for each different area during the year 2019: top-20 ranking areas.](image)

The area 5501 (Via Riva Reno 68 K043) achieved a total of 43005 occurrences. In particular, as a case study, we choose to focus on this area by evaluating the drivers’ daily behavior. Figure 4 shows the hourly statistic on parking meters amount reported in terms of sum (a), mean (b) and standard deviation (c). The highest number of amounts are present in the central hours of the day (from 9 am to 6 pm) while the lowest number of amounts are present during the night (from 3 am to 5 am).

Starting from these preliminary results, we measure if there is a daily effect inside the data, i.e. how much the data is correlated across different hours of the day (see Figure 6). We disclose how the parking meters amount during the night is negatively correlated with those during the day. In fact, an 11-hour time-series shift leads to a negative correlation up to 0.5, thus pointing out the opposite daily/night drivers’ behaviors and confirming how the drivers’ behaviors are mainly focused during the day (from 9 am to 6 pm).

**Forecasting**

Before training the ML model, we resample the dataset in an hourly time slot by computing the related mean of candidate observations. Thus each observation is related to a specific hour of a specific day. To stabilize the variance of the time series and to move close to a Gaussian-like distribution we perform a log-transformation. As case study, we focus on eight (see Table 1) of the most crowded areas reported in Figure 3.

![Figure 4: Hourly statistic on parking meters amount: focus on area 5501 Via Riva Reno 68 K043](image)
Figure 5 shows the forecasting results of the proposed ML model for 4 different areas: area 5501: Via Riva Reno 68 K043 (see Figure 5a), area 15056: Via dell’Ospedale (fr.maternità) K374 (see Figure 5b), area 13020: Viale Giurini G. Margherita K664 (see Figure 5c) and area 1042: Via Capo di Lucca, Via del Pallone K374 (see Figure 5d). The graphs refer to the forecast for the last fifteen days of the year 2019 and the first day of the year 2020 (next 24h forecast). The next 24h predicted amount accurately follows the ground-truth across different hours and days.

The quantitative results of the ML model forecasting are reported in Table 1 in terms of mean absolute error (MAE) and mean squared error (MSE). The minimum MAE was achieved for the 15057 area (Via dell’Ospedale (fr.maternità) K374), while the maximum MAE was achieved for the 5501 area (Via Riva Reno 68 K043). These results demonstrated how the proposed additive model is effective and accurate to provide the next 24-h forecast of parking meters amount, with an average MAE across to different areas up to 0.61 €.

Conclusions

Our work aims to propose a PRS based on the ML additive approach for forecasting the next 24-h parking occupancy in terms of parking meters amount. The obtained forecasting results suggest that the proposed approach is a viable solu-
<table>
<thead>
<tr>
<th>Area</th>
<th>Desc</th>
<th>MAE</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>5501</td>
<td>Via Riva Reno 68 K043</td>
<td>0.68</td>
<td>0.76</td>
</tr>
<tr>
<td>15056</td>
<td>Via dell’Ospedale (fr.maternità) K374</td>
<td>0.56</td>
<td>0.54</td>
</tr>
<tr>
<td>13020</td>
<td>Viale Giurini G. Margherita K664</td>
<td>0.55</td>
<td>0.53</td>
</tr>
<tr>
<td>1042</td>
<td>Via Capo di Lucca, Via del Pallone K374</td>
<td>0.63</td>
<td>0.64</td>
</tr>
<tr>
<td>1040</td>
<td>Via del Pallone</td>
<td>0.62</td>
<td>0.62</td>
</tr>
<tr>
<td>2001</td>
<td>Piazza del Baraccano</td>
<td>0.64</td>
<td>0.66</td>
</tr>
<tr>
<td>15057</td>
<td>Via dell’Ospedale (fr.maternità) K374</td>
<td>0.54</td>
<td>0.50</td>
</tr>
<tr>
<td>12083</td>
<td>Via Zanolini 12 K374</td>
<td>0.64</td>
<td>0.64</td>
</tr>
</tbody>
</table>

As a future direction step, we plan to convert the collected amount in terms of parking occupancy by taking into account the different occurrences and the corresponding hourly rate. Starting from this pilot study, we are currently integrating the proposed ML model in the PRS system where ad-hoc data visualization interface will be provided to the user. At the same time, we are measuring how the PRS system is able to generalize across different cities by also integrating additional data such as weather conditions and holiday calendar. Future work may be addressed to refine the PRS system to work on different spatial resolution (street, crossroads, etc...) and improve the dataset annotation by including also data acquired by mobile app.

Finally, another interesting future direction could be addressed to compare the proposed algorithm with respect to state of the art deep learning sequential model (i.e. long short term memory network) that are widely used for solving time series prediction task.

Acknowledgments
This work was supported within the research agreement between Università Politecnica delle Marche and Pluservice srl.

References