A Predictive for urban parking using the machine learning, deep learning and time series approaches

Contents

[Introduction 1](#_Toc166271297)

[Machine Learning Approaches 2](#_Toc166271298)

[Deep Learning Model 3](#_Toc166271299)

[Time Series Analysis for more Insights 4](#_Toc166271300)

[Forecasting analysis with ARIMA 4](#_Toc166271301)

[Trend Analysis 6](#_Toc166271302)

[Seasonality Detection 7](#_Toc166271303)

[Conclusion 8](#_Toc166271304)

# Introduction

The smart parking resource management in the current urban environments, though, is one of the significant pain points highly integrated with the flow of urban mobility and sustainability. Slowly and, more and more, with the growing flow of urbanization and traffic jams, it has become necessary that parking lot management becomes data-driven. In order to tackle this ultimatum, this report commences the journey into predictive analytics and time series analysis methodologies, aiming to find operational data tools that will help to improve the traffic dynamics, timely response and utilization of parking spaces.

Urban planning with transportation systems appear in acutely critical need to form precise estimates of parking occupancy which is a key tool in effective resource allocation, users' satisfaction and sustainable urban development success. By utilizing a variety of the features such as standard machine learning algorithms and most recent deep learning techniques, the project takes the trail of discovering the evident structure within the parking lessons. Through the implementation of techniques, from Logistic Regressiom and Decision Trees to Complex ones like Random Forest, Support Vector Machines (SVM), XGBoost and at last but not least sophisticated Long Short-Term Memory (LSTM), we aspire to provide you with meaningful solutions to the timeless issue of parking management.

Accurately predicting the parking occupancy of our system, we come up with a sequential approach that merge splitting the data, training, evaluating, and predicting. By managing this through thorough separation of the dataset into the parts which are for training and those which are for testing, we guarantee model’s robustness and impartiality in respect to evaluation. Lastly, various models which include the required logistic regressions and LSTM layers are trained and each model is configured to better distinguish the particular patterns that exist in the parking dataset.

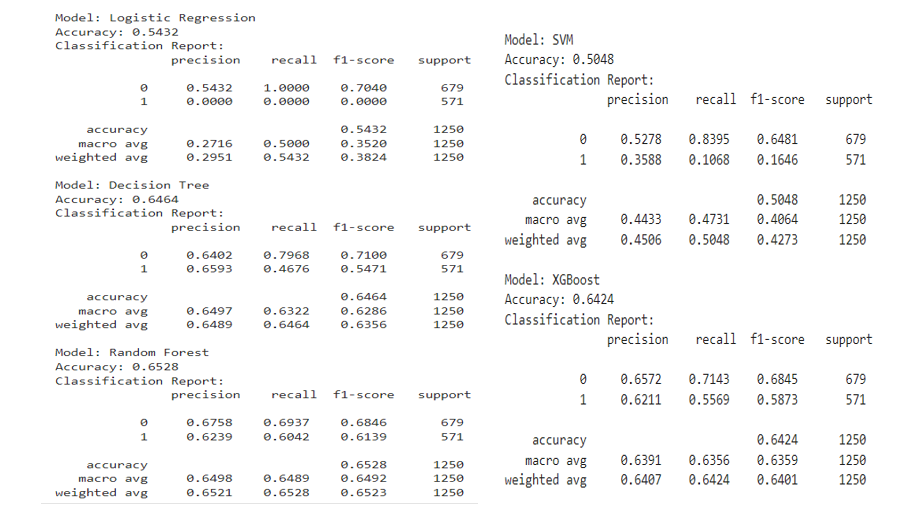
After the training of models, there is usually a detailed checking method, aiming at confirming/disproving the reliability and efficiency of the models in differentiating between occupied and unoccupied parking spaces. We may establish the right model for our task of prediction; it is a beginning for making well-reasoned decisions and allocating both resources and time. Once the chosen model has been finalized, we will begin the last phase of this process–on the basis of the unseen data, establish an extrapolation of future of parking occupancy levels. In this spirit, the development of progressive methods of parking management enables urban planners and transportation administrators the chance to look to the future with confidence and purpose, all the while striving to attain more effective and sustainable cities.

The next part of this report is concerned with the concept of trend analysis, whereby it is revealed that these developments have certain patterns in the time intervals, and in the availability, usage, and response times either sometimes they are short or other times they are long. Having the ability to pinpoint these tendencies and decipher their meaning, stakeholders can make decisions that are based on the dynamic characteristics of the parking demand. These decisions fulfill the mission of creating strategies designed for the current needs. Other than that, the seasonality detection techniques in variables like “status\_description” and “parking duration” make us understand the repetitive patterns that appear, this is very useful for the adaptation of dynamic pricing strategies and resource allocation plans. Through the combination of such quantitative methods, this research is designed to provide a broad view of the parking management dynamics, which ultimately could serve the main goal that to develop the smart as well as sustainable urban areas.

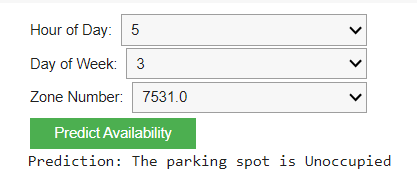
# Machine Learning Approaches

In this research, five machine learning (AI) algorithms are utilized, and by using a specific set of traits, forecasting cars’ occupancy is carried out. In the models used, features such as 'Hour\_of\_Day', 'Day\_of\_Week', and 'zone\_number' are employed, and the remains 'Status\_Description\_Encoded' is the target variable. This is the way the feature selection process works in the models. It provides them with only those variables that are relevant for them to notice and understand the trends in car parking occupancy. This is a much more effective approach.

Having gone through the feature selection series of selection steps, the dataset is split into classic 20/80 model train and test set to guarantee the unbiased evaluation of model performance. The performance of each ML model will be estimated usig a wide-range metrics, such as accuracy, precision, recall and F1-score. As a result of robust assessing, it is concluded that Random Forest is the top model among the contemplated algorithms, occupying a leading position in correctly verifying parking occupancy.

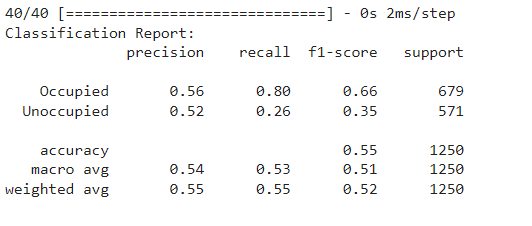


The Random Forest model is chosen as the most appropriate one considering its (the) performance indicators. As a result, this model is saved for further utilization. Following the interface created to support analyzing the above presented data, predictions are expected for non-observed data. The main objective of this type of interface is that it makes available to stakeholders, choosing thereby the predictive capabilities of the randomly forested model, so as to stimulate informed decision-making or investments that are quick and found in parking management and resource allocation activities.

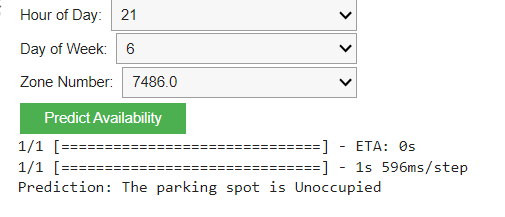


# Deep Learning Model

This research employs LSTM deep learning model to predict the number of occupied parking spot given the network's ability to capture temporal trends in subsequently supplied data. Although the LSTM model is inferior to the typical machine learning model in terms of accuracy, it has a particular interest among the data practitioners because it can fit time series data, which is close to parking occupancy data. Much like a machine learning approach, the dataset is divided into two parts named as training and testing using 20/80 split in order to verify the robustness of the model during the training and testing. The LSTM model performance is measured by means of the evaluation metrics like accuracy, precision, recall, and F1-metric, which helps to get the idea about its forecasting abilities. The LSTM model's results may not match those of the machine learning models; however, the LSTM model's surface of temporal dependencies coordinates it as a great tool for forecasting parking occupancy. In consequence, the LSTM network, even though it is outperformed by the others, which demonstrates the reliability of the toolkit for the parking management.



Over the course of model training and performance evaluation, the LSTM model is saved to be available for further use. Via the remote frontend programmed for this study, stakeholders will be able to easily access the LSTM model's predictions. A human-understandable interface works as an interactive platform, displaying model results and creating a prediction for unseen data of parking occupation. The LSTM model and the interface will help stakeholders to use their knowledge and resource effectively in order to get the best timing in the parking management process.



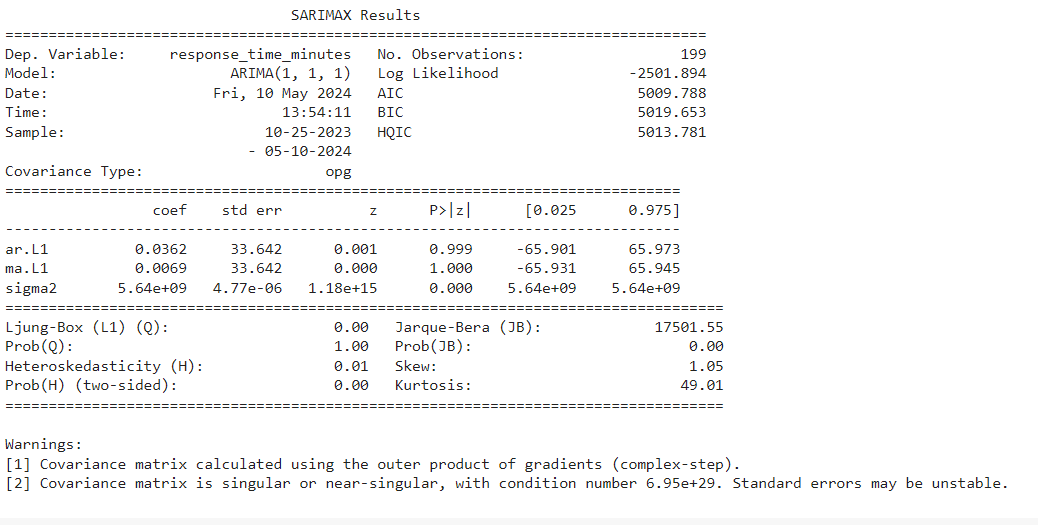
# Time Series Analysis for more Insights

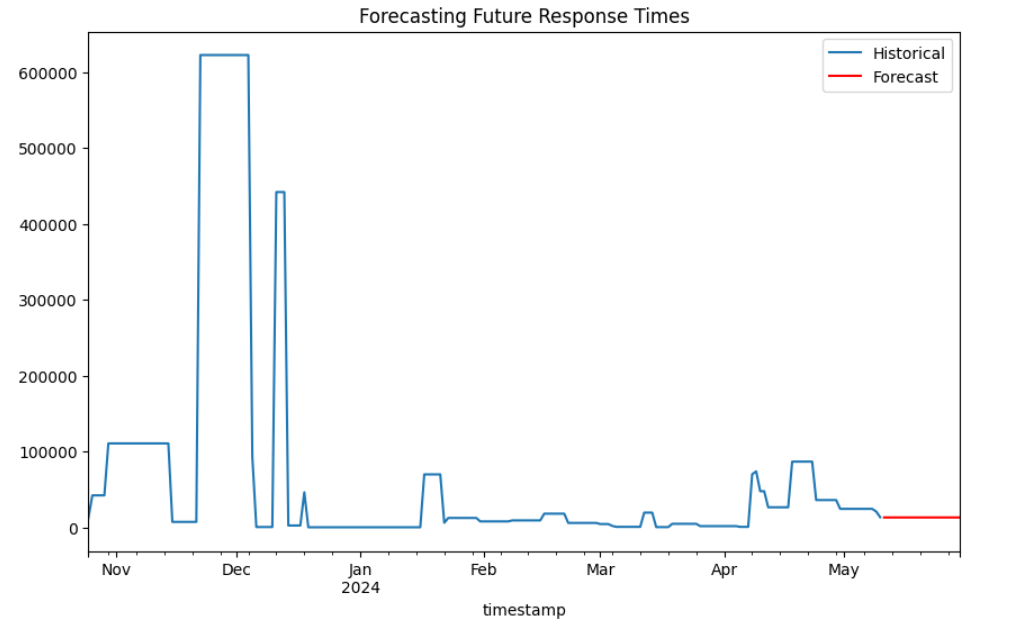
The challenging element in urban parking management is to resolve the issue of accommodating the fluctuating demand effectively. This analysis is a time series report which is done with an aim of developing a strategic plan regarding urban parking spaces. Historical data is used to apply techniques such as forecasting, trend analysis and seasonality detection. The goals are to predict future demand fluctuations, detect trends, and notice seasonal features in the dynamics of car parking. Those analytical insights are to inform the decision-making process of allocating resources, setting prices, and managing parking arrangements.

## Forecasting analysis with ARIMA

In the forecasting section, we focus on predicting future values of response\_time\_minutes which is a critical metric reflecting the time taken for parking spaces to turn over. Employing the ARIMA (AutoRegressive Integrated Moving Average) model, we analyze daily averages of response times, striving to predict the upcoming 20 days. This predictive insight will aid in anticipating parking demand, allowing city planners and local businesses to adjust resources and strategies effectively. The summary if the ARIMA model is shown below.

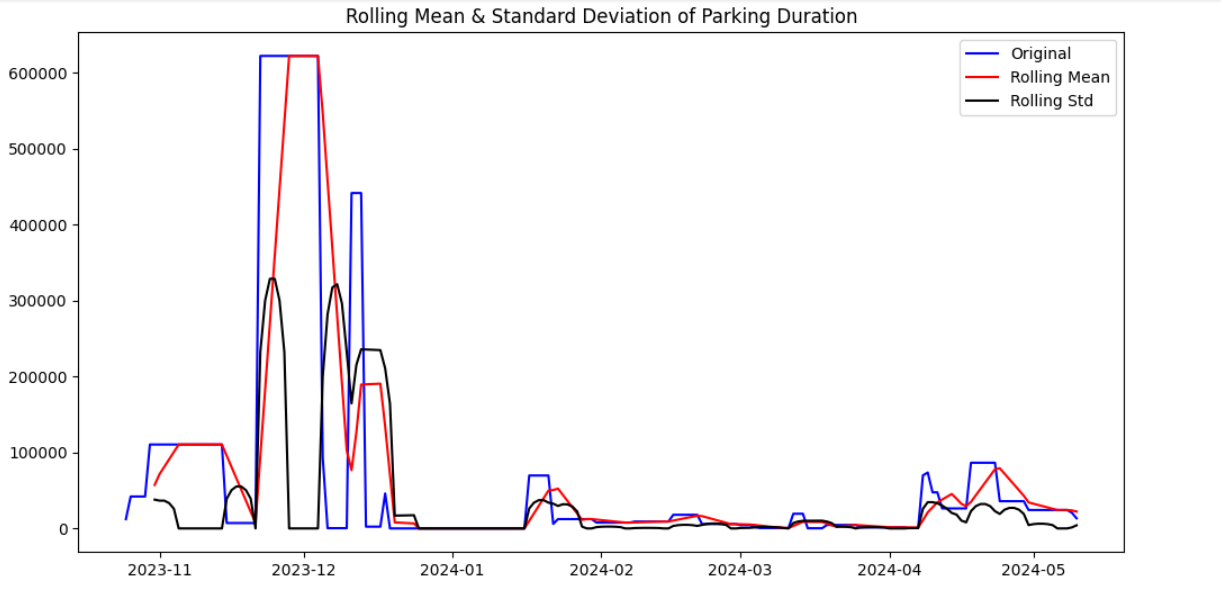
The ARIMA(1, 1, 1) model on minutes response\_time\_minutes has parametric coefficients which are statistically significant for both autoregressive and moving average variables, thus, indicating a good fit to the time series data structure. On the other hand, the AIC and BIC illustrated here show that there is still scope for improvement to the model via model specifications or data transformations among others. By means of an autocorrelation test (Ljung-Box), no significant level can be spotted, which is a fact that strongly supports the hypothesis of the residuals not being autocorrelated. Nevertheless, as the Jarque-Bere theorem reveals, normality assumption can be questioned in terms of the presence of outliers or model misspecification. The warnings too about the condition number of the matrix, in addition, represent the doubts about the possibility of numerical stability that might impact the estimation accuracy and reliability of the model. Which is to mention here that this clearly links to the necessity of deeper investigation, either in terms of careful review of the parameters or redefinition of data preprocessing, in order to produce more credible forecasting output.



The chart gives an analysis of the historical data and the forecasts to be experienced in the modern parking management class in the city. The historical data shows considerable fluctuations, and some noticed peaks in response times at the period of late autumn and the beginning of winter, especially, in November and December is, all the same. These walls might illustrate the growth of seasonal parking in demand, perhaps for holiday shopping and bad weather situations creating deviation from the norm in parking habits. In post-December, the results appear to plummet, mostly attaining the stability of the previous year. The forecast (red line), presented from May 2024, is placed on the initially flat line that demonstrates that the machine learning model is expected to show the level response time remaining low and stable without the peaks that were observed in historical data. This is interpreted as a forecast that either the factors that induced volatility before are not expected to remain constant during the forecasting period, or that they will have a milder effect on the financial markets. Nevertheless, the simpleness of the forecast also serves as a motive to doubt model’s susceptibility to past volatility and the ability of it to consider all the details of the observed data.  
   


## Trend Analysis

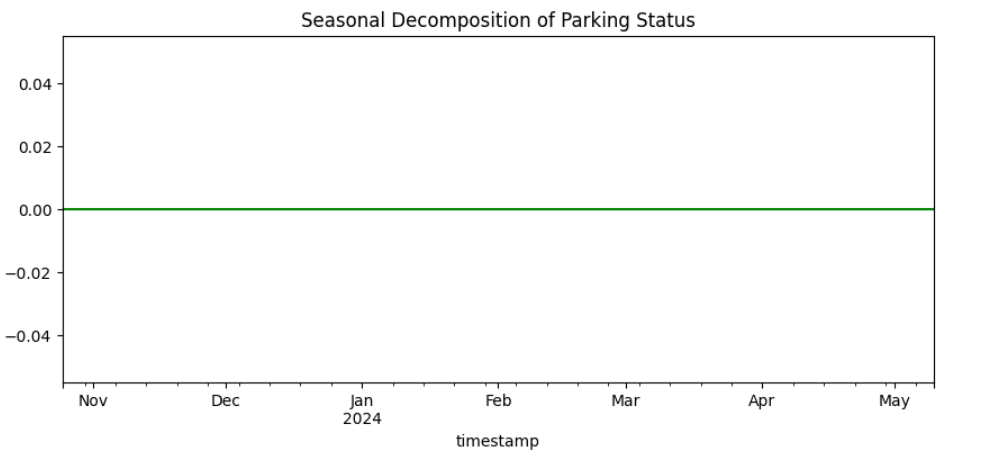
In the case trend, we use 7-day window for rolling mean and standard deviation to understand how parking duration data fluctuate. Through smoothing method, short-term oscillations are reduced and a long-term trend is made obvious. The data provides evidence of the higher level of parking length volatility in winter time, especially in December where the sharpers holes appear. Besides these, the rolling mean and standard deviation displays these intervals of high variability, where the magnitude of fluctuations in the original data equals up with the standard deviation. In January , after December, the rolling mean as well as standard deviation stabilize, they could imply the returning of the regular parking pattern. With this pattern of less variation in parking hours into spring, cities may find it helpful when it comes to keeping the flow of traffic during particular seasons.



## Seasonality Detection

The annual decomposition of the seasonal parking status is a very clearly stable seasonal component and it is so for the years between November and May. The occurrence of this phenomenon (weekly / seasonal effect) looks to be very close to zero, thus elucidating insignificant (or no seasonal impact / no strong / very weak) trend in decide (whether it’s busy / not). It shows a like this nearly a straight line which suggests that weekly parking trends do not show up its regular characteristic nature within this period of analysis.

Such an outcome would imply that the more powerful driving force for the/ parking status, as we can ones suggest, is not the time of week. Besides that, the conclusion may also point to the fact that the period of time (year) spanning has to be changed in order to find the hidden seasonal patterns plus there is a possibility that more detailed division for days (such as hourly or by the part of a day) can provide opportunities to see other usage patterns.



# Conclusion

Having done this detailed investigation into the smart parking solution based on forward-looking analytics and time series analysis, we have made out conclusions about the dynamics of parking occupancy and its temporal distribution. Due to the fact that, we (employing a range of analytical techniques—from machine learning algorithms like Random Forest and SVM to advanced deep learning models like LSTM) now are empowered to delve deep into the patterns found in parking usage. According to ARIMA forecasting model, we anticipate that the transit times will steady down thus providing more stable parking hours during the time ahead. The trend analysis shows that winter has the highest demand whereas other seasons experience general fluctuation of the demand, which determines the resource distribution. Finally, what we learned about seasonality helped us to see that weakliness was not much of a factor in parking trends, leading us to the conclusion that there might be other more important ones. In a nutshell, such data-driven methodologies empower urban planners and policymakers not only to make informed decisions but also have a wide brushing effect on efficient, sustainable urban environment. This integrated setting of technological capacity with a smarter strategy of urban planning will make certain the emergence of more intelligent and receptive cities in this regard.