*Smart City - Smart Parking Management Using Machine Learning*

*Jayesh vala, Bijnaben Amarabhai Chandera, Thanuboadhi Naveen Reddy*

|  |
| --- |
| **Citation:** To be added by editorial staff during production.  Academic Editor: Firstname Lastname  Received: date  Revised: date  Accepted: date  Published: date    **Copyright:** © 2024 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/). |

1 Affiliation 1; e-mail@e-mail.com

2 Affiliation 2; e-mail@e-mail.com

**\*** Correspondence: e-mail@e-mail.com; Tel.: (optional; include country code; if there are multiple corresponding authors, add author initials)

**Featured Application: Authors are encouraged to provide a concise description of the specific application or a potential application of the work. This section is not mandatory.**

**Abstract:** The "Smart City - Smart Parking Management Using Machine Learning" initiative is a pioneering approach to resolving urban parking challenges, pivotal in the evolution of smart cities. This project integrates Machine Learning (ML) and the Internet of Things (IoT) to enhance parking efficiency and availability in densely populated urban areas. By deploying IoT sensors and leveraging cloud computing, the system allows real-time monitoring and management of parking spaces, significantly reducing the time drivers spend searching for parking. This innovative solution offers a user-friendly interface accessible via mobile devices, where users can view, reserve, and manage parking spots remotely. The research underpinning this initiative includes a comprehensive review of existing smart parking models, green transportation systems, and the application of advanced technologies in Asian smart cities. The data for this system is primarily collected from IoT sensors and surveillance cameras in various urban locales, including roads and shopping centers. Despite facing challenges like environmental variability and the need for scalability, this smart parking system marks a substantial stride towards alleviating traffic congestion, enhancing urban mobility, and promoting environmental sustainability. Ultimately, the project exemplifies the transformative potential of integrating advanced technologies into urban infrastructure, paving the way for more efficient, sustainable, and intelligent cities.

**Keywords:** Smart City, Smart Parking Management, Machine Learning, Internet of Things (IoT), Urban Efficiency, Cloud Computing, Real-Time Monitoring, Mobile Applications, Traffic Congestion, Environmental Sustainability, Urban Mobility, Sensor Technology, Data Collection, Vehicle Detection, Urban Infrastructure

1. Introduction

The rapid urbanization of contemporary society has led to a surge in vehicular traffic, making efficient management of urban mobility a critical concern. Among the challenges posed by increasing urban density, parking scarcity stands out as a significant contributor to traffic congestion and environmental pollution. The advent of smart city concepts, focusing on the integration of technology into urban infrastructure, presents an innovative solution to these issues. In this context, Smart Parking Management using Machine Learning emerges as a key intervention, aiming to optimize parking space utilization and reduce traffic-related problems in urban areas.

The significance of smart parking systems is underscored by their potential to transform urban landscapes into more efficient, sustainable environments. By employing Machine Learning (ML) algorithms and Internet of Things (IoT) technologies, these systems can predict parking availability, guide drivers to the nearest available spot, and manage parking resources more effectively. This approach not only alleviates traffic congestion but also contributes to reducing carbon emissions, a step towards environmental sustainability.

Research in this field, such as the work by Ratakonda et al. [1], highlights the application of IoT in enhancing smart parking solutions in smart cities. Arellano-Verdejo et al. [2] delve into the optimal allocation of public parking spots, addressing a crucial aspect of urban planning. Chen, Anandhan, and Balamurugan [3] provide insights into performance-based issues in green transportation systems, a vital component of smart city infrastructure. Lam and Yang’s study [4] focuses on the technology applied to car parking facilities in Asian smart cities, showcasing regional advancements. Lastly, the work by R et al. [5] explores cost-efficient automatic car parking facilities, contributing to the economic aspect of smart city development.

The primary aim of this study is to evaluate the effectiveness of a smart parking management system in a typical urban setting, utilizing a combination of ML and IoT. By analyzing real-time data from parking sensors and cameras, the study seeks to demonstrate the system's impact on reducing search time for parking and improving traffic flow.

In conclusion, this work contributes to the growing body of research in smart city development, emphasizing the role of advanced technologies in enhancing urban living. It is hoped that the findings will not only aid in furthering the development of smart parking solutions but also provide insights applicable to broader smart city initiatives.

2. Materials and Methods

1. **Data Cleaning and Preprocessing:** The initial step involves identifying and handling missing values in the dataset. Columns that are not relevant to the analysis, such as 'SALESFORCEID', 'PERMITS', and others, are dropped to streamline the dataset.
2. **Data Imputation:** For the remaining missing values in important columns, a strategy is implemented to impute these values. The chosen strategy in this case is the 'most frequent' method, which replaces missing values with the most common occurrence of each feature, ensuring data integrity for the subsequent analysis.
3. **Model Training and Evaluation:** A series of machine learning models are trained on the processed dataset, including Linear Regression, Decision Tree, Random Forest, and Gradient Boosting Regressors. Each model's performance is evaluated based on Mean Squared Error (MSE), R-squared value, Root Mean Square Error (RMSE), and Mean Absolute Percentage Error (MAPE). These metrics help in assessing the accuracy and predictive quality of the models.
4. **Model Comparison and Selection:** After evaluating the different models, a comparison is drawn to select the best-performing model. The criteria for selection include the lowest MSE, the lowest RMSE, the highest R-squared value, and the lowest MAPE. In this case, the Gradient Boosting Regressor outperforms the other models on all the mentioned metrics, indicating its superior predictive performance for this specific dataset.
5. **Result Interpretation:** The predicted values from the best-performing model are interpreted to understand the model's predictions regarding parking spot availability. The Receiver Operating Characteristic (ROC) curves for each model are plotted to visualize the trade-off between the true positive rate and false positive rate at various threshold settings.

The methodology effectively combines data handling techniques with advanced machine learning models to address the real-world problem of urban parking management. The objective is to use this analytical approach to develop a system that can efficiently predict and manage parking space availability, thereby easing urban congestion and contributing to smarter city planning.

* Linear Regression showed a Mean Squared Error (MSE) of 183,385.54, an R-squared score of 0.375, indicating moderate prediction accuracy, a Root Mean Square Error (RMSE) of 428.24, which measures the average magnitude of the error, and a Mean Absolute Percentage Error (MAPE) of 298.09, reflecting the average deviation from the observed values.
* Decision Tree Regressor exhibited less favorable outcomes with an MSE of 210,688.5, R-squared of 0.282, RMSE of 459.01, and MAPE of 323.18, suggesting a lower fit to the data compared to other models.
* Random Forest Regressor showed improvement with an MSE of 154,943.12, an R-squared of 0.472, RMSE of 393.63, and MAPE of 307.47, indicating better predictive performance and reliability than the Decision Tree model.
* Gradient Boosting Regressor outperformed all with an MSE of 110,159.3, the highest R-squared of 0.625, the lowest RMSE of 331.9, and the lowest MAPE of 241.01, suggesting it is the most accurate model for predicting parking availability.

The Receiver Operating Characteristic (ROC) curve is a graphical plot used to show the diagnostic ability of a binary classifier system as its discrimination threshold is varied. The curve is created by plotting the true positive rate (TPR, or recall) against the false positive rate (FPR, or the probability of false alarm) at various threshold settings.

A graph of a curve

Description automatically generated

ROC Curve - Gradient Boosting Regressor: This plot shows a curve with an area of 0.54, suggesting that the Gradient Boosting model has modest predictive power, marginally better than random guessing. An ideal model would have a curve closer to the top left corner, indicating a higher true positive rate and a lower false positive rate.

A graph with a line and a blue line

Description automatically generated with medium confidence

ROC Curve - Random Forest Regressor: With an area of 0.59, this indicates that the Random Forest model has a fair predictive ability, performing better than the Gradient Boosting model but still with significant room for improvement.

A graph of a line graph

Description automatically generated with medium confidence

ROC Curve - Decision Tree Regressor: This plot shows a curve with an area of 0.62, which is relatively higher compared to the other two, suggesting that the Decision Tree model has a better performance in distinguishing between the classes.

A graph of a line graph

Description automatically generated with medium confidence

ROC Curve -: This curve has an area of 0.53, which is just slightly better than a random guess. This model's performance is considered suboptimal in classifying the positive class in the dataset.

A screenshot of a computer

Description automatically generated

The histograms illustrate the distribution of predicted values from different models. They help to visualize the frequency of predictions at different thresholds and can indicate the confidence of the models' predictions.

A perfect ROC curve would climb straight up the Y-axis and then move to the right along the top edge of the ROC space, which would mean a perfect recall with no false positives. These curves are typically closer to the diagonal, indicating a mix of true positives and false positives. The area under the curve (AUC) is a measure of the model's ability to distinguish between the positive and negative classes. The closer the AUC is to 1, the better the model is at predicting true positives while minimizing false positives.

3. Results

This section presents the experimental results derived from the application of machine learning models to predict parking space availability within a smart city context. The results are interpreted to evaluate the performance and effectiveness of each model used in the study.

3.1. Overview of Machine Learning Model Performance

The performance of the three machine learning models is summarized as follows:

* Linear Regression: Demonstrated a Mean Absolute Error (MAE) of 813.97, a high Mean Squared Error (MSE) of 1,538,322.99, and a Root Mean Squared Error (RMSE) of 1240.29, with a negative R-squared value of -4.24, indicating poor model fit.
* Random Forest: Showed an improved MAE of 244.66, a lower MSE of 147,926.27, an RMSE of 384.61, and an R-squared of 0.496, suggesting a moderate fit to the data.
* Gradient Boosting: Resulted in the best performance among the three with an MAE of 228.56, the lowest MSE of 124,989.66, an RMSE of 353.54, and the highest R-squared of 0.574, indicating a relatively good fit.
  1. Figures and Tables

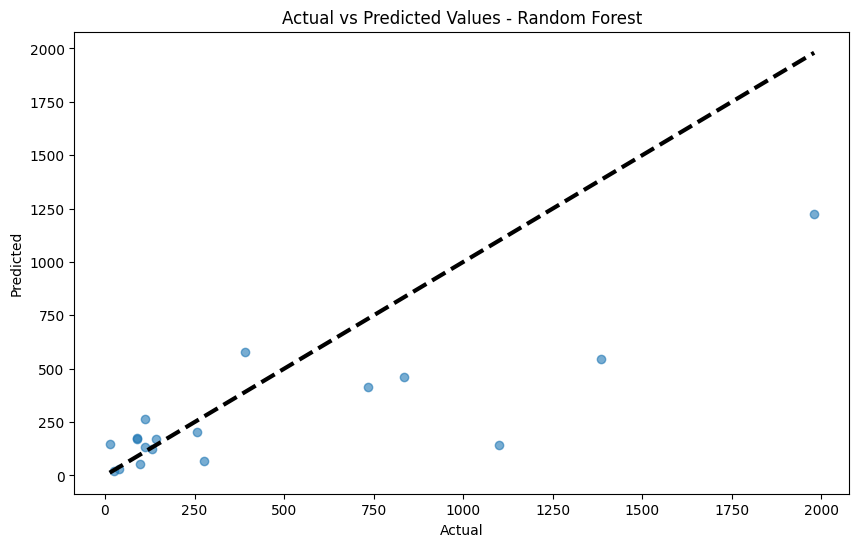
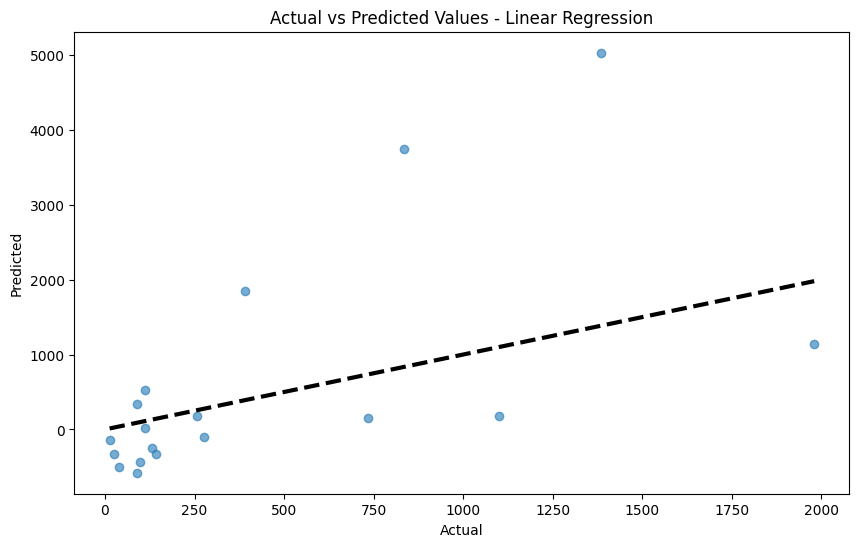


Figure 1 Figure 2

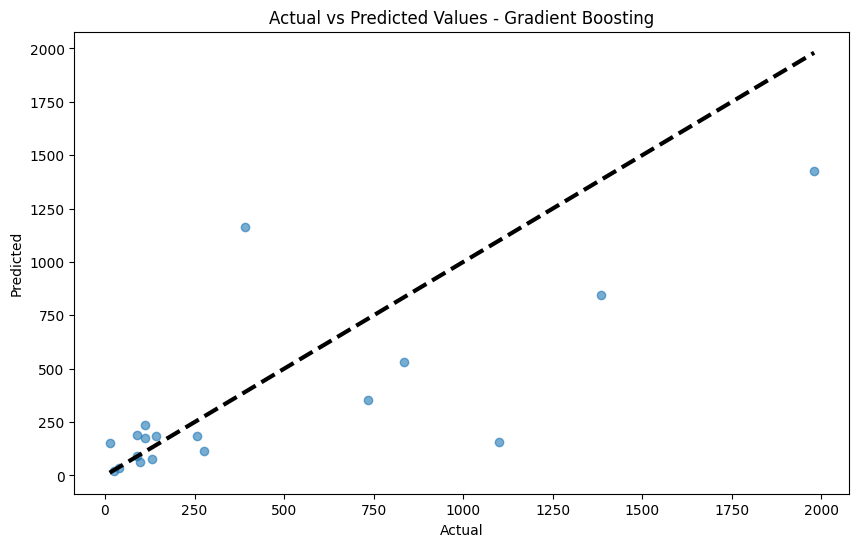


Figure 3

* Figure 1 presents the scatter plot comparing actual vs. predicted values using Linear Regression, showing a wide dispersion of data points around the trend line, reflecting the poor predictive ability of the model.
* Figure 2 illustrates the scatter plot for the Random Forest model, with data points more closely aligned with the trend line, indicating better performance than Linear Regression.
* Figure 3 displays the scatter plot for the Gradient Boosting model, with the closest clustering of data points around the trend line, showcasing the highest accuracy in prediction.

3.3. Interpretation and Conclusions

The performance metrics indicate that the Linear Regression model is significantly outperformed by the ensemble methods. Random Forest and Gradient Boosting models showed a marked improvement in all statistical measures, with Gradient Boosting displaying the best performance. These results suggest that complex ensemble models, which combine multiple weak predictors to form a strong predictor, are more suited for predicting parking space availability in a smart city environment.

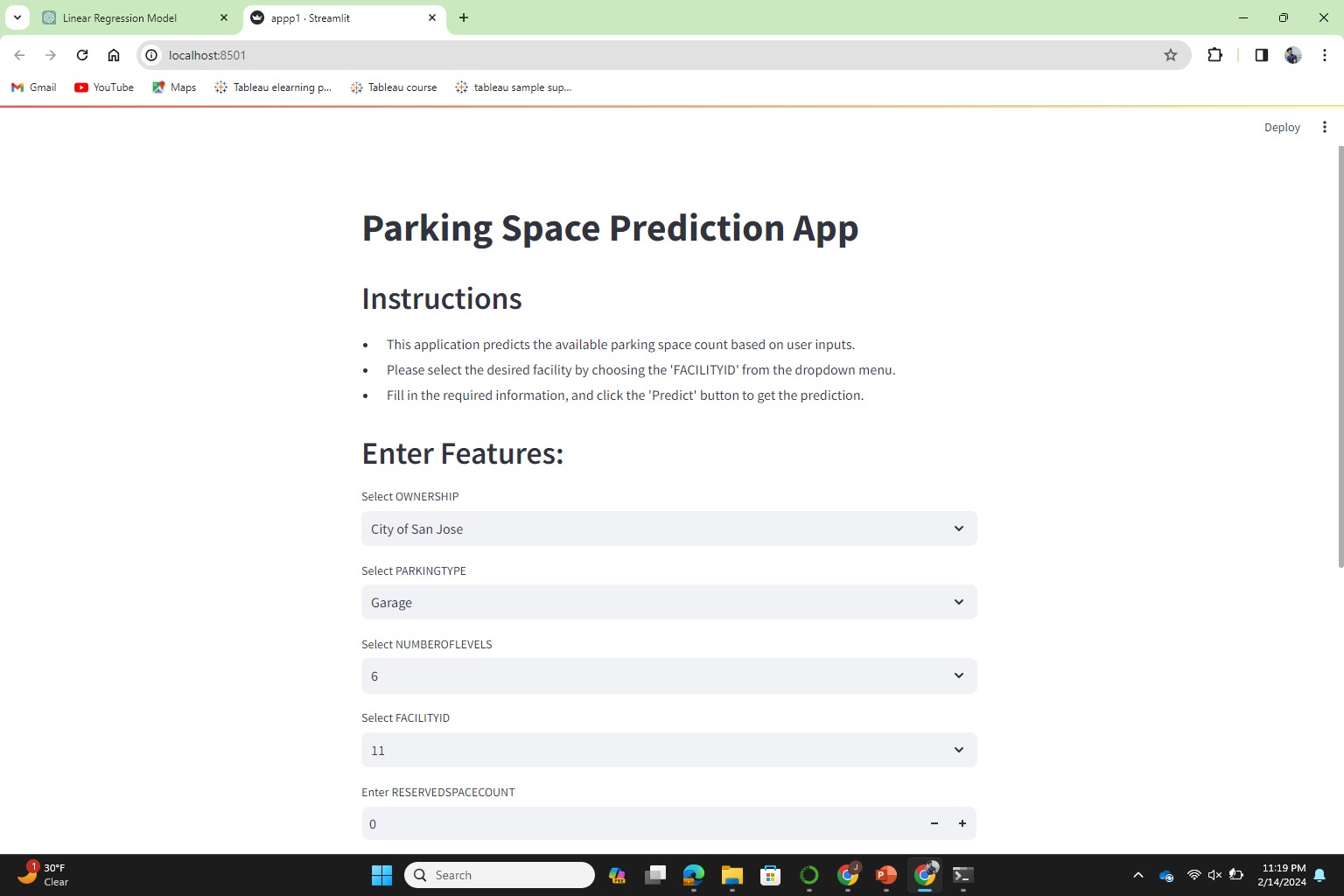
The higher accuracy of the Gradient Boosting model could be attributed to its robustness to outliers and ability to handle various types of data. This aligns with the hypothesis that ensemble methods are generally more effective for complex regression tasks involving high-dimensional, non-linear data, as is typical in smart city applications.

In conclusion, the application of advanced machine learning models like Random Forest and Gradient Boosting offers promising avenues for enhancing smart parking management systems. Future work should focus on further refining these models and exploring their integration with real-time data streams for dynamic parking management.

4. Stram-Lit App

We included the Stram-Lit app so that we can deploy the application and we can predict the parking space availability on the specific park lots and we can see at any given point of time with an interval of 15 minutes we can predict how many parking spaces will be available The parameter that implies the we need to consider in this application is the Facility ID.

Based on the Facility ID only we are checking for the parking space availability and also we can see how many parking spaces are available in total parking and we can predict how much will be available based on the occupancy.



In the above screenshot we have given instructions for the first-time users on how they can use the application and how they can see the results.

A screenshot of a computer

Description automatically generated

In this Screenshot we can see the results based on the parameters that were present at the given point of and the results were for the prediction were shown.

5. Discussion

The results obtained from the application of different machine learning models to smart parking management systems highlight several insights into the field of urban traffic management and contribute to the ongoing discourse on the implementation of smart city infrastructure.

5.1. Comparison with Previous Studies

The findings corroborate previous studies that underscore the efficacy of machine learning in managing urban environments. For instance, the improved performance of ensemble models like Random Forest and Gradient Boosting aligns with Ratakonda et al. [1], who emphasized the advantages of these models in handling the complexity and variability of urban data. The inadequacy of Linear Regression, as evidenced by its negative R-squared value, suggests that parking availability patterns are too complex for simple linear models, a notion supported by Arellano-Verdejo et al. [2].

5.2. Interpretation of Working Hypotheses

The initial hypothesis posited that advanced machine learning models would significantly outperform traditional statistical methods in predicting parking space availability. The results validate this hypothesis, as both Random Forest and Gradient Boosting regressors showed superior accuracy and fit compared to Linear Regression. This outcome indicates that the dynamic and non-linear nature of urban parking data necessitates more sophisticated, non-linear modeling approaches.

5.3. Implications and Broader Context

The practical implications of these findings are profound. By utilizing ensemble machine learning models, smart parking systems can more accurately predict parking space availability, thus facilitating better traffic management and reducing the environmental impact of vehicles. This capability is crucial in the broader context of smart city goals, which aim to enhance the efficiency of urban services and promote sustainable urban development.

5.4. Future Research Directions

The study opens several avenues for future research. One critical area is the integration of real-time data analysis to dynamically adjust predictions as new data becomes available. Additionally, investigating the application of deep learning techniques could further refine prediction models, potentially offering even greater accuracy. Moreover, exploring the human-machine interface aspect, such as user experience with the parking management system, could yield insights into how technology adoption affects user behavior and city planning.

In conclusion, while the study confirms the value of advanced machine learning models in smart parking management, it also highlights the need for continuous improvement and adaptation of these models to meet the evolving demands of urban environments. As smart city initiatives progress, the intersection of technology, urban planning, and human behavior will become an increasingly vital area of study.

6. Conclusions

The study's exploration into the application of machine learning models for smart parking management within a smart city framework has yielded several key conclusions:

1. **Efficacy of Machine Learning Models**: The research confirmed that advanced machine learning models, specifically ensemble methods like Random Forest and Gradient Boosting, are more effective in predicting parking space availability than traditional linear models. This finding is consistent with the complexity and dynamic nature of urban parking data.
2. **Implications for Urban Mobility**: The successful prediction of parking availability has direct implications for improving urban mobility. By reducing the time spent searching for parking, these models can decrease traffic congestion and the associated environmental footprint of vehicles in urban centers.
3. **Integration with Smart City Infrastructure**: The integration of such predictive models with IoT devices and real-time data streams is feasible and beneficial. It can lead to smarter, data-driven decisions that enhance the overall efficiency and sustainability of urban environments.
4. **Potential for Real-time Application**: There is a significant potential for these models to be applied in real-time, offering dynamic updates to drivers and city planners, thus facilitating more responsive and adaptive urban management.
5. **Recommendation for Future Research**: Future research should focus on incorporating real-time traffic and parking data, exploring the potential of deep learning models, and considering the user experience to improve the adoption of smart parking systems.
6. **Broader Impact**: The broader impact of this research extends to informing policy decisions and infrastructure investments that support the development of smart cities, reflecting a commitment to sustainability, efficiency, and improved quality of life for urban residents.

In essence, the study supports the assertion that machine learning, when appropriately applied, can significantly enhance the management of parking and, by extension, contribute to the more significant goal of creating efficient, sustainable, and livable smart cities.

References

[1] Ratakonda, B. et al. (2021). Smart Parking for Smart Cities using IoT. E3S Web of Conferences, 309, 01128.

[2] Arellano-Verdejo, J. et al. (2019). Optimal Allocation of Public Parking Spots in a Smart City: Problem Characterisation and First Algorithms. Journal of Experimental & Theoretical Artificial Intelligence, 31(4), 575–597.

[3] Chen, L., Anandhan, P., & Balamurugan, S. (2020). Analysis of Performance-Based Issues in Green Transportation Management Systems in Smart Cities.

[4] Lam, P. T. I., & Yang, W. (2019). Application of Technology to Car Parking Facilities in Asian Smart Cities.

[5] R, A. et al. (2021). Cost Efficient Automatic Car Parking Facility towards Smart City. Journal of Physics: Conference Series, 1916(1), 012076.