IEEE Access
Multidisciplinary · Rapid Review · Open Access Journal

Received 16 May 2024, accepted 5 June 2024, date of publication 7 June 2024, date of current version 17 June 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3411391

RESEARCH ARTICLE

Enhancing Urban Parking Efficiency Through Machine Learning Model Integration

SAI SNEHA CHANNAMALLU[1], SHARARЕH KERMANSHACHI[1], JAY MICHAEL ROSENBERGER[2], AND APURVA PAMIDIMUKKALA[1]

[1]Department of Civil Engineering, The University of Texas at Arlington, Arlington, TX 76019, USA
[2]Department of Industrial Engineering, The University of Texas at Arlington, Arlington, TX 76019, USA

Corresponding author: Sharareh Kermanshachi (sharareh.kermanshachi@uta.edu)

ABSTRACT An increase in vehicular traffic and a scarcity of parking spaces are creating significant challenges for urban parking management. This study aims to tackle these issues that escalate congestion and pollution and decrease urban productivity, by utilizing machine learning models to accurately predict parking space availability and categorize occupancy levels. It employs a dataset from a college campus garage collected from January 2022 to June 2023 and analyzes the performance of random forest, decision tree, linear regression, and support vector models by comparing them, using multiple evaluation metrics. The results revealed that the random forest model was the most reliable, as it demonstrated strong performance in both the regression and classification analyses and was adept at estimating the exact number of available parking spaces. A concurrent classification analysis that categorized parking occupancy into different levels proved valuable for enhancing the quality of communication and decision-making. An analysis of the importance of various features clearly highlighted the influence of the day of the week on parking demand and patterns; the impact of seasonality on the volume of parking usage; and the time of day, which plays a crucial role in determining parking behavior. The research will benefit urban planners, facility managers, and policymakers by providing them with insights and tools that will enhance the urban parking experience and address the complex challenges of modern urban environments.

INDEX TERMS Parking management, predictive modeling, random forest model, decision tree, support vector machine, linear regression.

I. INTRODUCTION

It is becoming progressively difficult to manage parking in urban settings due to the ever-increasing number of vehicles on the road and the constrained availability of parking spaces. Vehicle ownership has surged with economic growth and urban expansion, intensifying the imbalance between the availability and demand for parking [1]. Drivers typically spend from 3.5 to 14 minutes looking for parking spaces, depending on location and time, which induces stress and delays, increases pollution, and wastes fuel [2]. Reducing the amount of time spent searching for parking spots can diminish the circulation of vehicles in parking lots, thereby lowering traffic congestion and noise pollution around the entrances [3], [4], [5].

The inefficiency in parking searches stems from the absence of real-time information on space availability [6], [7], [8]. Providing drivers with access to real-time parking maps that reveal available spaces via their vehicle's navigation systems [9], [10] would not only benefit drivers but also aid parking facility operators and urban planners in managing their space more effectively, based on anticipated demand.

In the past, methods for predicting parking space occupancy were primarily based on static guidelines and heuristic approaches, which lacked accuracy and flexibility. These methods typically depended on direct observation, a narrow scope of past data, or elementary statistical techniques, which are not adequate for the intricate nature of city parking dynamics. However, the emergence of advanced technologies

The associate editor coordinating the review of this manuscript and approving it for publication was Yiqi Liu.

for data collection, the accessibility to wide-ranging parking data sets, and the incorporation of machine learning into these processes have marked a significant move towards approaches centered around data. According to [10] and [11], this shift has profound implications. Integrating machine learning into parking management has significantly improved the accuracy of forecasting parking demand, maximized the efficiency of space usage, and provided motorists with reliable information. This, in turn, reduces traffic congestion, enhances the satisfaction of the users, and boosts the overall operational efficacy [12], [13].

A. BENEFITS OF PARKING AVAILABILITY PREDICTION MODELS

Research indicates that the availability of parking spaces significantly influences drivers' decisions regarding where to park [14], and a lack of information about available parking can lead to increased time spent searching, cruising, and queuing, which is also costly. For example, in the US, drivers annually spend about 17 hours searching for parking, incurring a cost of around $345 per person [15]; in the U.K. and Germany, these figures are even higher. Conversely, drivers who have access to information on parking availability are less frustrated and are 45% more likely to make effective parking decisions [16], which reduces traffic and energy consumption and enhances transportation management [17], [18], [19].

Intelligent parking management systems and related research have shown that providing parking availability information enriches the customer's experience and enables online parking reservations [20], [21]. An example is SFpark, which was initiated in San Francisco in April 2011 and effectively manages on- and off-street parking availability, using intelligent parking meters that adjust pricing based on factors like location, time, and day and aim to keep about 15% of spaces vacant on each city block [22]. The system is equipped with sensors that track space availability and allow users to view prices via SFpark.org and mobile apps [23]. SFpark has shown notable results: a 4% reduction in metered parking rates and 12% in city garages, a 43% decrease in time spent searching for parking, and a 30% reduction in daily vehicle miles traveled, leading to safer streets, less congestion, and reduced neighborhood pollution. Additionally, SFpark has positively impacted local businesses, which have experienced a more than 35% increase in sales tax revenue compared to less than 20% in other city areas [24]. SFpark's achievements underscore the potential of predicting parking occupancy to tackle urban congestion and promote a more efficient, sustainable transportation system. Its use of sophisticated technology to predict parking availability and adjust prices has effectively diminished congestion and enhanced the user's experience [25], [26].

TABLE 1. Summary of previous research in the field.

Models	Results	Study
Nine different machine learning models including random forest, extra tree algorithm	The Random Forest algorithm stands out as the optimal machine learning model due to its efficiency and quick execution time, making it well-suited for accurately predicting large-scale, long-term parking space availability.	[27]
Diverse array of Machine Learning and Deep Learning techniques	Interestingly, it was noted that statistical models achieved performance levels comparable to those of more complex models.	[30]
Linear regression model artificial neural network (ANN) frameworks	A specific ANN model was suggested due to its high R-squared value of 0.846.	[31]
Logistic Regression. Decision Tree, K-Nearest Neighbors (KNN), Complement Naive Bayes	By representing parking space availability as a binary variable (0 for available spaces, 1 for unavailable), experimental results have shown promising outcomes	[29]
Random forest, decision tree, support vector regression (SVR)	The random forest model surpassed both the decision tree and SVR in terms of accuracy. Nonetheless, the decision tree demonstrated greater computational efficiency than the other models.	[28]
Genetic algorithm and SVR	Concluded that machine learning models accurately predict car park occupancy rates in smart cities	[32]
Decision tree and random forest	Decision tree-based techniques are more suitable in accurately predicting parking occupancy.	[33]
Decision tree and SVR	Highlighted the capability of decision tree models to precisely predict parking space availability.	[34]
Gradient boosting decision tree algorithms	Demonstrated the effectiveness of decision tree methods in predicting short-term parking space occupancy.	[11]
KNN, SVR, Random Forest	Concluded that the SVR model was the most accurate of those tested	[35]
Decision tree and random forest	Combining decision tree algorithms with genetic algorithms yielded dependable forecasts of parking space availability.	[12]
Support vector machine (SVM), K-nearest neighbor (KNN), random forest	Highlighted the significance of factoring in the type and size of parking to ensure precise parking occupancy forecasts and determined that the SVM algorithm was more precise than other techniques	[36]
SVM	Combined image processing algorithms with machine learning techniques and concluded that the experimental results demonstrated accuracy and effectiveness in real-world scenarios	[37]
Random forest, decision tree, SVR, linear regression	SVR showed superior recall performance, while linear regression exhibited the least mean absolute error.	[38]
	The random forest and decision tree models consistently outperformed the SVR and linear regression models in terms of accuracy and F1-score. The random forest model excelled in both precision and recall.	[39-45]
Decision trees, SVM, random forest	Machine learning models effectively predict parking space availability, provide valuable information to drivers, and contribute to improved parking management in urban areas.	[46]
Random forest, SVR, and neural networks (NN)	The deep NN model demonstrates accurate parking occupancy prediction and achieves high evaluation metrics by leveraging deep learning techniques.	[47]
Decision tree, gradient boosting, SVR, random forest, and NN	The random forest and gradient boosting models yielded the lowest MAE and RMSE values, indicating higher accuracy in their predictions.	[48]

B. PROBLEM STATEMENT AND OBJECTIVE

While existing literature extensively explores the application of machine learning models in urban parking management, it fails to comprehensively investigate and promote understanding of the full spectrum of urban parking challenges.

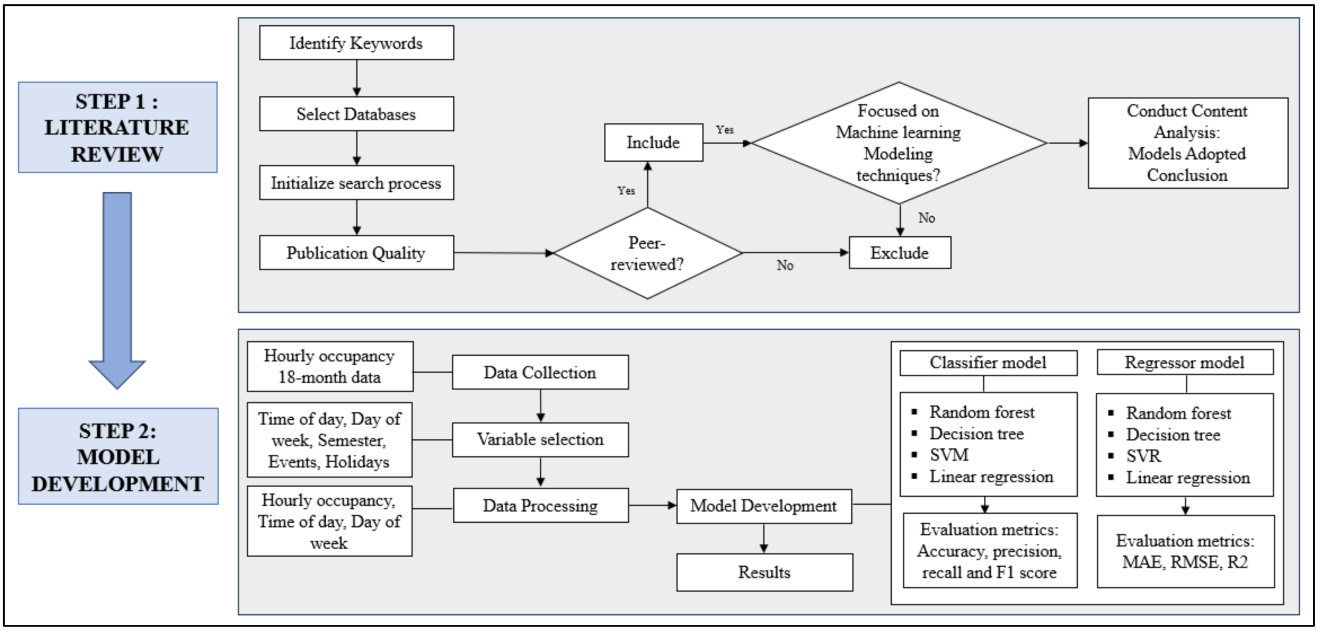


FIGURE 1. Research methodology.

Current studies primarily focus on the technical aspects of model accuracy and efficiency, but often overlook the practical integration of these models into real-world urban parking systems. In addition, there is a lack of research on the adaptability of the models to dynamic urban environments, where parking demands and patterns can vary significantly due to a variety of factors. This study aims to address these shortcomings by developing an intelligent model that can accurately forecast parking availability and examining the operational integration and adaptability of machine learning models in diverse urban scenarios. It seeks to bridge the gap between theoretical model accuracy and practical applicability, ensuring that the developed solutions are robust, scalable, and directly beneficial to urban parking management systems. The objectives of the study are to: (1) assess and compare machine learning models that can precisely predict the availability of urban parking spaces, (2) investigate the dual functionality of the chosen model in providing both specific predictions and categorical insights for real-time parking management, and (3) analyze the factors that affect parking demand and behavior. The findings from this study are of tangible relevance to the operations of parking management systems in urban environments, offering assistance in the development of effective and enduring parking strategies.

II. LITERATURE REVIEW

The complexity of urban environments has escalated the need for sophisticated parking management solutions, and a significant body of research has addressed this urban challenge by focusing on the application of machine learning models. Studies such as those by [27] and [28] have been instrumental in exploring the effectiveness of various machine learning algorithms in the context of urban parking, and the integration of machine learning in urban parking solutions, as highlighted by [29], reflects a growing trend towards the adoption of data-driven strategies in urban planning. Table 1 presents a summary of the literature on this subject.

The consistent effectiveness of random forest and decision trees in predicting parking availability and occupancy is a notable trend in the literature. Works by [12], [27], [28], and [41] and others have shown that these models outperform others in terms of accuracy, precision, and F1-score, indicating their reliability for urban parking management systems. Decision trees are noted for their computational efficiency in rapid processing, which is an important attribute for real-time parking management applications, as highlighted in the study by [28] and makes them a practical choice for optimizing urban parking. Comparing the performance of various models suggests that the efficacy of a particular model can be context-specific and influenced by factors such as the type of parking and scale of data. Studies by [36] and [38] indicate variations in model performance based on these parameters and emphasize the need for a tailored approach to model selection and implementation that considers the specific characteristics of the parking environment.

III. METHODOLOGY

The methodology depicted in figure 1 is a structured approach to conducting research on machine learning models for parking prediction. The research methodology began with performing a literature review, in which keywords related to parking predictions and machine learning were identified. Appropriate databases were then searched for scholarly articles, with a focus on publication quality and peer-reviewed status. Articles that focused on machine learning modeling techniques for parking prediction were included, while others were excluded. The last step of the literature review was a content analysis of the selected models.

The next step was to collect data and select the variables for development of the machine learning models. Data processing was conducted to prepare the data, then was used to develop two types of models: classifiers and regressors. Evaluation metrics were used to assess the models' performance, and the results obtained, and conclusions drawn from the model development phase were discussed.

A. DATA COLLECTION

The dataset consisted of 13,104 entries that encompassed the date and hourly occupancy data from January 2022 through June 2023; the hourly occupancy percentages are depicted in figure 2. For assessing previous occupancy levels, a retrospective analysis method was utilized, incorporating a range of inputs that provided the model with information on occupancy rates from an hour before predictions were made.

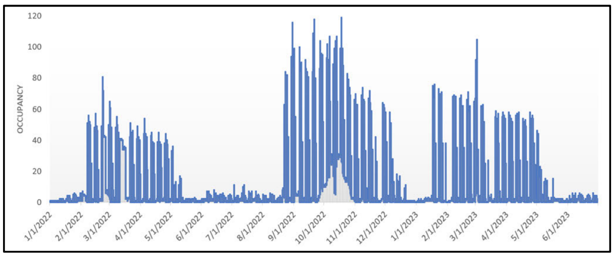


FIGURE 2. Data set.

B. VARIABLE SELECTIONS

Key features such as the day of the week, time of day, and contextual factors like academic semesters, holidays, and special events were chosen for their potential influence on parking occupancy. Figure 3 illustrates the occupancy variations according to these selected variables.

The selected features, including the day of the week, time of day, academic semesters, holidays, and special events, were chosen based on their demonstrated influence on parking occupancy patterns in the specific context of college campus parking. For instance, the day of the week was deemed crucial due to its significant impact on parking demand, with weekdays experiencing higher occupancy levels than weekends, driven by class schedules and staff activity. Similarly, the time of day emerged as a key predictor, reflecting temporal variations in parking demand throughout the day, with notable differences observed between peak commuting hours and off-peak periods. Academic semesters were included to capture seasonal variations in parking patterns, with distinct occupancy levels observed during fall, spring, and summer terms. e.g., occupancy during summer semesters is generally lower than during the fall and spring semesters. Additionally, holidays and special events were incorporated to account for transient disruptions in parking demand, ensuring the models can adapt to anomalous fluctuations and produce accurate predictions. The study also accounted for semester breaks, as parking occupancy during these periods is usually lower due to the absence of classes and fewer students on campus. Additionally, the analysis included significant events such as exams, popular sports events, and concerts, recognizing that these events can cause notable fluctuations in standard parking patterns that often result in occupancy surges.

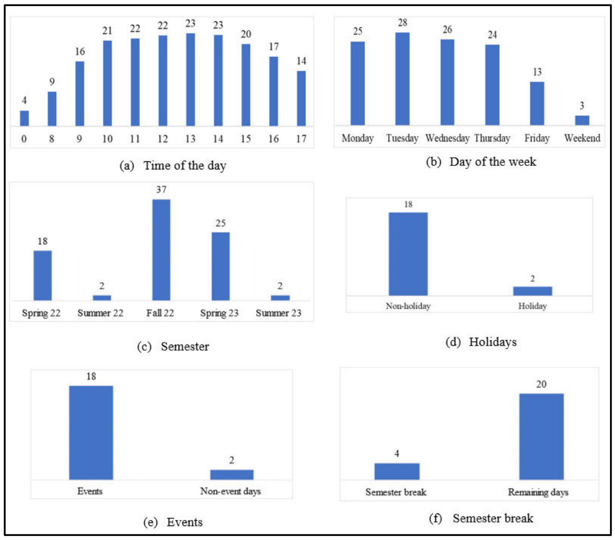


FIGURE 3. Variations in average hourly occupancy.

Overall, the choice of features was guided by empirical evidence and domain expertise, aiming to capture the multifaceted dynamics of parking occupancy and enhance the models' predictive accuracy and robustness.

C. DATA PREPROCESSING

This study analyzed hourly occupancy data, both as a continuous and as a categorical variable. The data was classified into three categories (low, medium, and high), based on tercile distributions, with specific thresholds defined for each group to enhance clarity and improve prediction accuracy. The analysis primarily focused on the typical hours of a college's operation, which is from 8 am to 5 pm, and data outside these hours were averaged and aligned with the 0-hour mark. This approach was adopted to fine-tune the analysis by emphasizing the most pertinent timeframes. The data for Saturdays and Sundays were combined and treated as a single "weekend" category, simplifying the analysis and acknowledging the different parking patterns that are generally seen on these days.

D. PREDICTION METHODS

Four models were chosen as contenders for the model that most accurately predicts parking space occupancy at a college campus: random forest, SVM, linear regression, and decision trees. All four possess superior predictive abilities, but the random forest model is particularly distinguished for its robust management of complex variable interrelations and its effectiveness in reducing overfitting, a frequent issue in prediction models. The decision tree model demonstrates straightforwardness and easy interpretability and offers a clear decision-making path based on input features, which makes making it user friendly for non-technical stakeholders and ideal in situations where it's important to understand how the model arrives at decisions. Linear regression is valued for the uncomplicated manner in which it models linear relationships between variables and outcomes, proving effective when the linear model assumption holds true. It is also simple to use and understand, making it suitable for initial analysis or linearly related data. SVM is known for its superior handling of nonlinear relationships and is essential in complex situations like parking occupancy forecasts. Its ability to navigate nonlinearity makes it a viable option for intricate predictor-target relations. Table 2 presents a summary of the parameters utilized for each model, along with the values obtained through optimization for each parameter.

Python libraries were employed to operationalize these models. To better address the time-sequential characteristics of the dataset, akin to time series data, we implemented k-fold cross validation with $k = 10$ folds. This method not only maintains the temporal sequence of the data during training but also allows for a more comprehensive assessment across different segments of the dataset. By partitioning the dataset sequentially into 10 folds, we ensure that each validation set consists of data points that immediately follow their corresponding training set, preserving the chronological integrity and avoiding leakage of future information. This approach is critical for evaluating the models' ability to adapt and respond to cyclical patterns, providing a robust validation of predictive performance over time.

TABLE 2. Model parameters.

Models	Optimal parameters
Random forest	no. of estimators: 100, maximum depth: 15, minimum samples split: 4, minimum samples leaf: 2
Decision tree	maximum depth: 56, minimum samples split = 16, minimum samples leaf = 2,
SVM	regularization parameter (c): 1.0, kernel: radial basis function (rbf), tol: 0.1, degree: 3
Linear regression	Default parameters.

E. EVALUATION METRICS

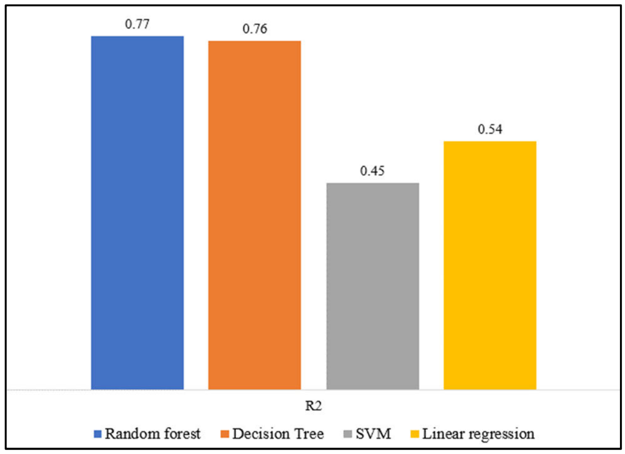
The models' ability to accurately predict parking occupancy was analyzed by performing a regression analysis, using the mean absolute error (MAE), root mean square error (RMSE), and R-squared (R^2) as primary metrics. Absolute measures, RMSE and MAE, were compared to identify the model with the lowest average error in predictions; R^2 was used to quantify the percentage of variance in the dependent variable that the model explains. In classification analysis, precision calculates the ratio of true positives to all positive predictions, recall evaluates the model's capability to identify all actual positives correctly, and the F1 score acts as a balanced metric between precision and recall, essentially serving as their harmonic mean.

IV. RESULTS

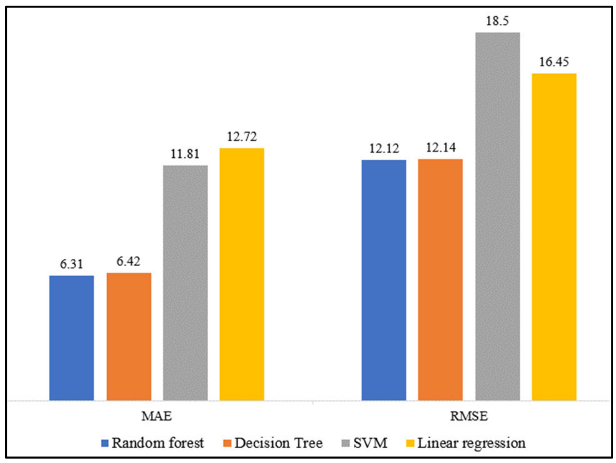
The comparative analysis of the random forest, decision tree, linear regression, and SVR models highlighted each model's specific advantages and limitations in forecasting parking space availability in a college campus garage.

A. REGRESSION ANALYSIS

An analysis of the machine learning models that used hourly occupancy as a continuous variable revealed the differences in the performance and suitability of each model. Figure 4 presents the average R^2 score, MAE and RMSE across all folds in the cross-validation of the models before pre-processing the hourly occupancy.



(a) Average R2 score



(b) Average MAE and RMSE scores

FIGURE 4. Results of models.

The random forest model stood out for its excellence in predicting hourly parking occupancy, as it exhibited an MAE of 6.31, an RMSE of 12.12, and an impressive R^2 score of 0.77. The decision tree model was closely aligned with the random forest model in performance, with an MAE of 6.42, a similar RMSE, and an R^2 score of 0.76. The linear regression model showed a weaker performance with an MAE of 12.72, an RMSE of 16.45, and an R^2 score of 0.45. The SVR model displayed an MAE of 11.81, an RMSE of 18.5, and an R^2 score of 0.54.

Figure 5 depicts the importance of features and shows that the latter half of the workweek, from Wednesday to Friday, was given greater importance in the regression analysis. This suggests that these days are more indicative of high parking usage and may reflect a pattern in which parking demand peaks during the mid-to-late workweek due to increased activities or consistent scheduling patterns. The seasons also exhibited a clear impact on parking usage, with fall emerging as the most influential season, closely followed by spring and winter. This trend could be attributed to the start of the academic year and activities that typically increase parking demand during these times. The significant influence of fall might be due to back-to-school periods and new academic sessions, which typically see a surge in campus activity and, consequently, parking demand.

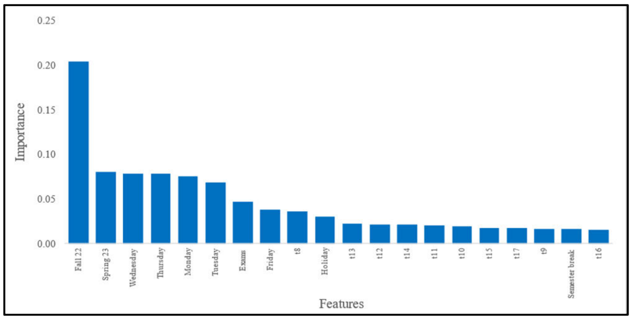


FIGURE 5. Feature importance score derived from regression analysis.

The time-of-day features, denoted as f1 through f17, show a range of importance scores that reflect the varying parking demands throughout the day. These may correspond to typical daily routines such as morning arrivals, lunchtime exits, and evening departures, which all influence parking space turnover and availability. Holidays and weekends were revealed to be of moderate importance, which indicates their role in the predictability of parking usage and also suggests a more consistent parking pattern on these days compared to regular weekdays.

B. CLASSIFICATION ANALYSIS

The bar chart presented in figure 6 shows comparisons of the performances of four machine learning models (random forest, linear regression, SVM, and decision tree), based on four key metrics: precision, recall, F1 score, and accuracy. In terms of Precision, the Random Forest model achieved the highest score with 0.99, followed by Linear Regression at 0.97, SVM at 0.95, and the Decision Tree model at 0.71. For Recall, Random Forest led with 0.98, closely followed by Linear Regression and SVM scoring 0.95 and 0.93 respectively, and the Decision Tree at 0.72. In measuring F1 Score, the Random Forest again topped the chart with a score of 0.99, with Linear Regression and SVM scoring slightly lower at 0.96, and the Decision Tree trailing at 0.72. Finally, in terms of Accuracy, Random Forest excelled with a score of 0.99 followed by linear regression at 0.96, SVM also performed robustly with a score of 0.94, and the Decision Tree showed an accuracy of 0.75

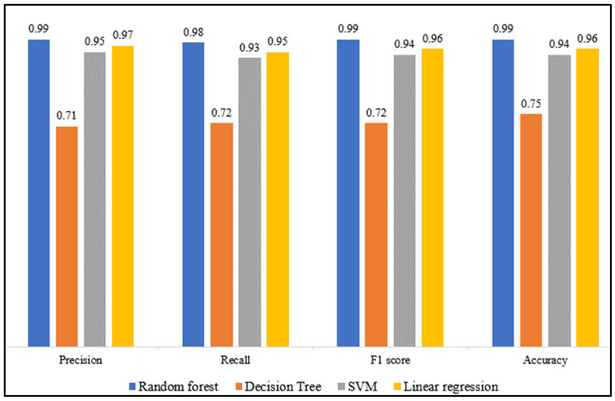


FIGURE 6. Results of prediction models.

Figure 7 shows that Monday is the most influential factor in the classification analysis, underscoring the impact of the beginning of the workweek on parking trends. This is closely followed by temporal features labeled f15, f17, and Wednesday, all indicating critical periods within the weekly cycle that shape parking behaviors. Thursday, f13, and the spring season are shown to be of moderate importance, suggesting a tangible but less dominant influence on parking dynamics.

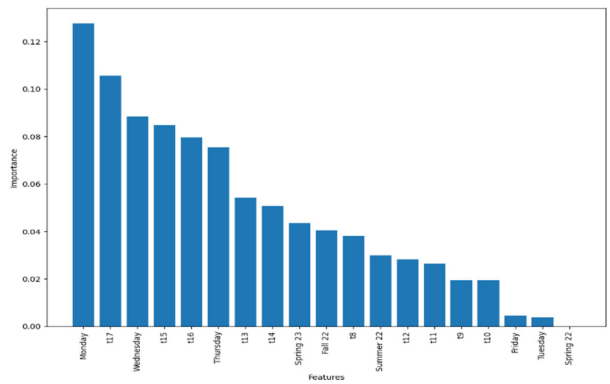


FIGURE 7. Feature importance scores derived from classification analysis.

The lesser importance attributed to the fall and summer seasons, as well as specific time slots denoted by f8 through f10 and days like Friday, holidays, and Tuesday reveals that these factors have less impact on the classification of parking patterns. This information is indicative of a lower variability in parking behavior or a less acute effect on the differentiation of occupancy levels during these periods.

V. DISCUSSION

The results of this study show that the random forest model excelled in capturing the complexities of parking occupancy data, with high accuracy, precision, and an ability to explain a significant portion of the variances in the dependent variable. The decision tree algorithm also showcased notable effectiveness, albeit with a marginally higher MAE compared to the random forest, indicating its reliable predictive power and its capability to generate accurate forecasts. The linear regression model's notably higher MAE and RMSE values indicate less precise predictions and greater variance in errors, making it less effective for capturing the variability of parking occupancy data. The SVR model demonstrated moderate precision but showed significant variance in its predictive errors and struggled to effectively capture underlying patterns in the data. Random forest and linear regression had superior performance across all metrics in the classification analysis, which indicated that the dataset was highly reliable. The SVM also showed strong performance, particularly in recall and its F1 score, and the decision tree model, despite its lower performance in precision and accuracy, still presents a respectable performance, particularly in recall. It may, however, benefit from further fine-tuning to improve its precision and overall accuracy.

Several important insights that can optimize urban parking emerged from comparing the scores assigned to various features by the classification and regression analyses. Days of the week were shown to play a crucial role in shaping parking behaviors across both methodologies, which suggests that daily cycles are essential determinants in predicting both the number of parking spaces needed (quantitative aspect) and the types of parking behavior (qualitative aspect). Seasonality was revealed to have a pronounced impact in regression analysis, highlighting its strong influence on the actual use of parking spaces throughout different times of the year. Its role in classification is more moderate, however, indicating that while seasonal changes affect parking usage, they do not drastically alter the types of parking behaviors. The time of day shows a greater impact in classification analysis, indicating its importance in identifying various parking behaviors during specific times, such as peak and off-peak hours, yet this factor has a lesser influence in regression analysis, implying that while significant, the time of day does not heavily sway the overall quantity of parking needed. This comprehensive evaluation highlights the multifaceted nature of urban parking patterns and illustrates how different factors influence parking behaviors through the lenses of classification and regression analyses.

A. PRACTICAL INTEGRATION CHALLENGES AND LIMITATIONS WITHIN REAL-WORLD URBAN PARKING SYSTEMS

Expanding the understanding of practical integration challenges and limitations within real-world urban parking systems is crucial for contextualizing the applicability of the study's findings. Implementing machine learning models for forecasting parking space availability in operational contexts presents a myriad of challenges and complexities. Firstly, the availability and quality of data pose significant hurdles. Real-time data collection from parking facilities may be hindered by technological and logistical constraints, impacting the reliability and timeliness of model predictions. Additionally, deploying infrastructure such as sensors and IoT devices entails considerable costs and maintenance efforts. Addressing compatibility issues and ensuring seamless integration with existing urban infrastructure and management systems is essential for practical implementation. Moreover, ethical considerations surrounding data privacy and user consent necessitate careful attention to regulatory compliance and community engagement efforts. Furthermore, the ongoing maintenance and adaptation of machine learning models, coupled with user acceptance and behavior dynamics, require sustained efforts from interdisciplinary stakeholders. Collaborative initiatives involving researchers, urban planners, policymakers, technology providers, and community representatives are essential for overcoming these challenges and realizing the potential benefits of machine learning-based parking prediction systems in diverse urban environments.

B. COMPARISON WITH EXISTING STUDIES

The findings of this study, particularly the superiority of the random forest model, are aligned with the findings of [27], [39], [40], and [41], who also reported higher accuracy and F1-scores for this model compared to other machine learning models. Reference [36] highlighted the superior accuracy of the SVM algorithm in certain contexts, however, suggesting that the choice of the model may depend on the specific characteristics of the parking environment, such as type and scale. The literature review reveals that the choice of the model can be highly context specific. Studies by [36] and [38] suggest that the performance of machine learning models varies based on factors such as parking type and data scale, which aligns with our study's observation on the importance to model accuracy of considering seasonal variations and the time-of-day. Our research supports the notion that while models like random forest are generally effective, their performance can be enhanced by tailoring them to the specific characteristics of the parking environment.

This research expands upon existing studies, which predominantly focus on individual machine learning models [21], [22], [23], by assessing a broader array of models specifically tailored for predicting parking occupancy. Unlike most previous work that uses datasets spanning less than six months [28], [29], [30], [31], [32], [33], [34], [35], [36], [37], [38], [39], [40], [41], [42], [43], [44], [45], [46], [47], [48], this study employs a more comprehensive 18-month dataset, enabling a deeper analysis of trends and patterns that enhances the robustness of its conclusions and predictions. Furthermore, whereas studies such as [24] generally target large-scale predictions, this research zeroes in on urban parking space availability, incorporating factors beyond mere execution time. In contrast to [26], which explores various indicators within parking phenomena, this study focuses explicitly on parking occupancy in urban settings, addressing a critical aspect of urban management essential for reducing congestion and optimizing space utilization. Additionally, this study distinguishes itself by emphasizing the practical implications for real-time parking management, highlighting the effectiveness of the random forest model in delivering detailed predictions and categorizations, thereby making substantial contributions to effective parking management strategies [28].

VI. CONCLUSION

Parking management in urban areas has become increasingly challenging due to the rising number of vehicles and limited number of parking spaces. This study employed a comprehensive analysis of machine learning techniques to predict parking occupancy in an urban setting, based on a dataset of hourly occupancy data collected over a period of 18 months that provided a solid foundation for evaluating the effectiveness of the various models. A comparative analysis was performed of four models (SVM, random forest, decision tree, and linear regression) to assess the performance of each, utilizing metrics such as MAE, RMSE, and R2 for regression analysis, and accuracy, precision, recall, and F1 score for classification analysis. This study not only identifies the most effective models for predicting parking occupancy but also illuminates the various factors that influence parking patterns.

The in-depth evaluation of both regression and classification models revealed the exceptional performance of the random forest model, whose perfect or near-perfect scores in classification metrics and remarkable proficiency in regression analysis unequivocally establishes it as the optimal choice for predicting parking occupancy. This dual strength allows for both detailed, number-specific predictions and categorizations that serve the diverse needs of real-time parking management, enabling users to make quick assessments and facility managers to develop effective strategic plans and allocate their resources wisely. The study's exploration into feature importance scores for urban parking optimization revealed that urban parking patterns are influenced by the days of the week, seasonality, and the time of day. The days of the week impact both parking demand and behaviors; seasonality impacts the volume of parking usage, especially in quantitative assessments; and the time of day plays a crucial role in discerning various parking patterns in qualitative analysis. This underlines the intricacy of urban parking management and emphasizes the necessity for a comprehensive strategy that accommodates daily, seasonal, and hourly nuances in parking trends.

While meticulous attention was given to this detailed study on predicting parking occupancy, it's crucial to recognize the study's limitations. While the methodology and findings offer valuable insights applicable to a broad range of parking scenarios, the data set may not apply to other venues. It should be noted, however, that the success of the random forest model in this environment underscores its potential adaptability to other contexts, albeit with careful consideration of their unique characteristics. This study lays a strong foundation for future research, which could further enhance the utility of our findings by incorporating real-time data, dynamically adjusting models, and implementing these predictive models in diverse parking management systems.

ACKNOWLEDGMENT

The authors gratefully acknowledge the support and generosity of the North Central Texas Council of Governments (NCTCOG), without which the present study could not have been completed.

DATA AVAILABILITY

All data generated or analyzed during the study are available from the corresponding author by request.

DECLARATION OF COMPETING INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this article.

REFERENCES

[1] Y. Yang, S. Sun, Z. Qiao, and H. Zhu, "Parking services in smart cities: A comprehensive survey," *IEEE Access*, vol. 7, pp. 64271–64292, 2019.

[2] Y. Zheng, S. Rajasegarar, and C. Leckie, "Parking availability prediction for sensor-enabled car parks in smart cities," in *Proc. IEEE 10th Int. Conf. Intell. Sensors, Sensor Netw. Inf. Process. (ISSNIP)*, Apr. 2015, pp. 1–6.

[3] F. Caicedo, C. Blazquez, and P. Miranda, "Prediction of parking space availability in real time," *Expert Syst. Appl.*, vol. 39, no. 8, pp. 7281–7290, Jun. 2012.

[4] S. S. Channamallu, S. Kermanshachi, J. M. Rosenberger, and A. Pamidimukkala, "Parking occupancy prediction and analysis—A comprehensive study," *Transp. Res. Proc.*, vol. 73, pp. 297–304, Jan. 2023.

[5] A. O. Kotb, Y.-C. Shen, and Y. Huang, "Smart parking guidance, monitoring and reservations: A review," *IEEE Intell. Transp. Syst. Mag.*, vol. 9, no. 2, pp. 6–16, Summer 2017.

[6] S. Yang, W. Ma, X. Pi, and S. Qian, "A deep learning approach to real-time parking occupancy prediction in transportation networks incorporating multiple spatio-temporal data sources," *Transp. Res. C, Emerg. Technol.*, vol. 107, pp. 248–265, Oct. 2019.

[7] A. Pamidimukkala, S. Kermanshachi, J. M. Rosenberger, and G. Hladik, "Barriers to adoption of electric vehicles in Texas," *Environ. Sci. Pollut. Res.*, vol. 31, no. 11, pp. 16735–16745, Feb. 2024, doi: 10.1007/s11356-024-32337-7.

[8] H. Huang, G. Gartner, J. M. Krisp, M. Raubal, and N. Van de Weghe, "Location based services: Ongoing evolution and research agenda," *J. Location Based Services*, vol. 12, no. 2, pp. 63–93, Apr. 2018, doi: 10.1080/17489725.2018.1508763.

[9] M. Sester, "Analysis of mobility data—A focus on mobile mapping systems," *Geo-spatial Inf. Sci.*, vol. 23, no. 1, pp. 68–74, Jan. 2020, doi: 10.1080/10095020.2020.1730713.

[10] S. S. Channamallu, V. K. Padavala, S. Kermanshachi, J. M. Rosenberger, and A. Pamidimukkala, "Examining parking occupancy prediction models: A comparative analysis," *Transp. Res. Proc.*, vol. 73, pp. 281–288, 2023.

[11] Z. Liu, Y. Li, Y. Liu, and H. Cai, "Parking occupancy prediction based on gradient boosting decision tree," in *Proc. IEEE 2nd Int. Conf. Inf. Syst. Comput. Aided Educ. (ICISCAE)*, Sep. 2019, pp. 358–362.

[12] Y. Sun, Y. Wang, L. Chen, and Y. Gao, "Short-term parking occupancy prediction based on decision tree and random forest," in *Proc. 3rd Int. Conf. Ind. Bus. Eng.*, 2019, pp. 135–139.

[13] R. K. Patel, S. S. Channamallu, M. A. Khan, S. Kermanshachi, and A. Pamidimukkala, "An exploratory analysis of temporal and spatial patterns of autonomous vehicle collisions," *Public Works Manage. Policy*, 2023.

[14] S. Hendricks and M. Outwater, "Demand forecasting model for park-and-ride lots in king county, Washington," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 1623, no. 1, pp. 80–87, Jan. 1998.

[15] INRIX Research. (2017). *Searching for Parking Costs Americans $73 Billion a Year*. Accessed: Aug. 24, 2023. [Online]. Available: http://inrix.com/press-releases/parking-pain-us

[16] F. Caicedo, F. Robuste, and A. Lopez-Pita, "Parking management and modeling of car park patron behavior in underground facilities," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 1956, no. 1, pp. 60–67, Jan. 2006.

[17] D. C. Shoup, "Cruising for parking," *Transp. Policy*, vol. 13, no. 6, pp. 479–486, Nov. 2006.

[18] S. S. Channamallu, S. Kermanshachi, J. M. Rosenberger, and A. Pamidimukkala, "A review of smart parking systems," *Transp. Res. Proc.*, vol. 73, pp. 289–296, Jan. 2023.

[19] T. Lin, H. Rivano, and F. Le Mouël, "A survey of smart parking solutions," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 12, pp. 3229–3253, Dec. 2017.

[20] *Feasibility Study of Real Time Parking Information at Metrorail Parking Facilities—Virginia Stations*, Washington Metrop. Area Transit Authority, Wilbur Smith Associates, Bengaluru, India, 2009.

[21] N. Adepu, S. Kermanshachi, A. Pamidimukkala, and E. Nwakpuda, "An analytical approach to understanding construction cost overruns during COVID-19," *Smart Sustain. Built Environ.*, Apr. 2024.

[22] R. Gordon. (2010). *High-Tech Parking Meters Premiere in S.F.* SFGate. [Online]. Available: https://www.sfgate.com/bayarea/article/High-tech-parking-meters-premiere-in-S-F-3257977.php

[23] M. Richtel. (2011). *Now, to Find a Parking Spot, Drivers Look on Their Phones*. Accessed: May 7, 2023. [Online]. Available: https://www.nytimes.com/2011/05/08/technology/08parking.html

[24] SFMTA. (2017). *San Francisco Adopts Demand-Responsive Pricing Program to Make Parking Easier*. [Online]. Available: https://www.sfmta.com/blog/san-francisco-adopts-demand-responsive-pricing-program-make-parking-easier

[25] C. K. B. Krishnamurthy and N. S. Ngo, "Parking, transit and traffic: Evidence from SFpark," CERE Work. Paper, p. 6, Jun. 2018, vol. 2018. [Online]. Available: https://ssrn.com/abstract=3245146

[26] A. Pamidimukkala, S. Kermanshachi, J. M. Rosenberger, and G. Hladik, "Barriers and motivators to the adoption of electric vehicles: A global review," *Green Energy Intell. Transp.*, vol. 3, no. 2, Apr. 2024, Art. no. 100153.

[27] A. N. H. Soumana, M. B. Salah, S. Idbraim, and A. Boulouz, "Machine learning models in the large-scale prediction of parking space availability for sustainable cities," *Endorsed Trans. Internet Things*, vol. 10, no. 2024, Nov. 2023.

[28] M. E. Yalcin and M. Zeydan, "Comparative analysis of machine learning algorithms for parking occupancy prediction," *Expert Syst. Appl.*, vol. 63, pp. 249–259, Nov. 2016.

[29] Y. Filali, H. Errousso, N. Aghbalou, E. A. A. Alaoui, and M. A. Sabri, "Real-time parking availability classification on a large-area scale," in *Innovations in Smart Cities Applications* (Lecture Notes in Networks and Systems), vol. 906, M. B. Ahmed, A. A. Boudhir, R. E. Meouche, and İ. R. Karaş, Eds. Cham, Switzerland: Springer, 2023, doi: 10.1007/978-3-031-53824-7_20.

[30] S. Fiorini, M. Ciavotta, C. Liberto, and G. Valenti, "On-street parking prediction: A comparative study," in *Proc. IEEE Int. Conf. Big Data (BigData)*, Sorrento, Italy, Dec. 2023, pp. 282–291, doi: 10.1109/bigdata59044.2023.10386302.

[31] S. Paudel, M. Vechione, and O. Gurbuz, "Predicting university campus parking demand using machine learning models," *Transp. Res. Rec.*, vol. 2678, no. 6, pp. 14–26, 2023.

[32] Z. Song, J. Zhang, W. Zhu, and X. Xi, "The vector matching method in geomagnetic aiding navigation," *Sensors*, vol. 16, no. 7, p. 1120, Jul. 2016, doi: 10.3390/s16071120.

[33] J. Wu, B. Yang, and D. Zhu, "A comparative study of machine learning algorithms for parking occupancy prediction," *IEEE Access*, vol. 8, pp. 104431–104439, 2020.

[34] K. Dey and B. Nath, "Prediction of parking occupancy using decision tree and support vector machine," in *Proc. 3rd Int. Conf. Comput. Methodol. Commun. (ICCMC)*, 2019, pp. 131–134.

[35] V. Paidi, "Short-term prediction of parking availability in an open parking loT," *Int. J. Transp. Sci. Technol.*, vol. 8, no. 2, pp. 157–167, 2019, doi: 10.1515/jisys-2022-0039.

[36] Z. Zhao, Y. Zhang, and Y. Zhang, "A comparative study of parking occupancy prediction methods considering parking type and parking scale," *J. Adv. Transp.*, vol. 2020, Feb. 2020, Art. no. 5624586, doi: 10.1155/2020/5624586.

[37] S. Y. Bae, C. G. Kim, and J. K. Park, "Parking occupancy estimation based on image processing and machine learning techniques," *Sustainability*, vol. 12, no. 19, p. 8171, 2020.

[38] T. Anwar, S. Asadullah, M. Azzam, and K. Abbas, "Comparison of different machine learning models for parking occupancy prediction," in *Proc. Int. Conf. Frontiers Inf. Technol. (FIT)*, 2019, pp. 59–64.

[39] S. Farooq, W. Khaksar, M. A. Ullah, and I. Yaqoob, "Comparative analysis of machine learning algorithms for parking occupancy prediction," *J. Ambient Intell. Hum. Comput.*, vol. 10, no. 4, pp. 1343–1355, 2019.

[40] Q. He, Z. Li, C. Huang, M. Zhang, and D. Li, "Comparison of machine learning algorithms for parking occupancy prediction in urban areas," *J. Adv. Transp.*, vol. 2020, pp. 1–12, Apr. 2020.

[41] F. Wu, Z. Du, J. Zhang, S. Zhang, and Y. Wang, "Parking occupancy prediction based on decision tree and random forest," in *Proc. IEEE Int. Conf. Real-Time Comput. Robot. (RCAR)*, Jul. 2017, pp. 234–238.

[42] J. H. Kim and J. Y. Kim, "A comparative analysis of machine learning algorithms for parking occupancy prediction: Focusing on different time horizons," *Sustainability*, vol. 13, no. 10, 2021, Art. no. 5557.

[43] K. Zhang, S. Zheng, S. Zhao, and P. Li, "Comparative study of machine learning algorithms for parking occupancy prediction: A case study in Beijing," *Sustain. Cities Soc.*, vol. 75, Dec. 2021, Art. no. 103287.

[44] W. Shi, X. Chen, Y. Zhang, R. Guo, and Y. Xiong, "Comparative analysis of machine learning algorithms for parking occupancy prediction in smart city," *Wireless Commun. Mobile Comput.*, vol. 2021, Mar. 2021, Art. no. 5551553.

[45] S. Srinivasan, R. Soundararajan, K. V. Chalapathy, and R. Balasubramanian, "Comparative analysis of machine learning techniques for parking occupancy prediction," *J. Ambient Intell. Hum. Comput.*, vol. 12, no. 5, pp. 5789–5804, 2021.

[46] H. Brozova and M. Ruzicka, "The prediction of parking space availability," *Transport*, vol. 36, no. 3, pp. 302–311, 2021, doi: 10.3846/transport.2020.14016.

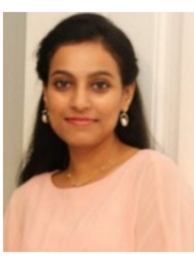
[47] S. Hasan, N. M. M. Nawi, and A. M. Zin, "Parking occupancy prediction using machine learning and sensor data: A comparative study," *J. Adv. Res. Dynam. Control Syst.*, vol. 13, no. 3, pp. 1052–1061, 2021.

[48] M. E. Basiri, I. Lozano-Márquez, and A. Fernández-Caballero, "A comparative study of machine learning techniques for parking occupancy prediction," *Sensors*, vol. 21, no. 4, 2021, Art. no. 1393.

[49] L. Gao, W. Fan, Z. Hu, and W. Jian, "Prediction of vacant parking spaces in multiple parking lots: A DWT-ConvGRU-BRC model," *Appl. Sci.*, vol. 13, no. 6, p. 3791, Mar. 2023.

[50] G. Li and X. Zhong, "Parking demand forecasting based on improved complete ensemble empirical mode decomposition and GRU model," *Eng. Appl. Artif. Intell.*, vol. 119, Mar. 2023, Art. no. 105717.

[51] K. Shang, Z. Wan, Y. Zhang, Z. Cui, Z. Zhang, C. Jiang, and F. Zhang, "Intelligent short-term multiscale prediction of parking space availability using an attention-enhanced temporal convolutional network," *ISPRS Int. J. Geo-Inf.*, vol. 12, no. 5, p. 208, May 2023.



SAI SNEHA CHANNAMALLU received the B.E. degree in civil engineering from the National Institute of Technology, Warangal, India, in 2016. She is currently pursuing the Ph.D. degree and a Graduate Research Assistant with The University of Texas at Arlington. Her research interests include autonomous vehicles, efficiency of parking systems, risk analysis, and cost optimization.

SHARAREH KERMANSHACHI received the bachelor's degree in architectural engineering, the master's degree in civil engineering, the Master of Business Administration (M.B.A.) degree, and the Ph.D. degree in civil engineering from Texas A&M University. She is currently a Professor with the Department of Civil Engineering, the Director of the Resilient Infrastructure and Sustainable Environment (RISE) Lab, and the Technology Transfer Director of the Center for Transportation, Equity, Decisions and Dollars (CTEDD), The University of Texas at Arlington. Her research interests include resilient infrastructures, post-disaster reconstruction, autonomous vehicles, probabilistic risk analysis, and construction performance optimization.

JAY MICHAEL ROSENBERGER received the B.S. degree in mathematics from Harvey Mudd College, the M.S. degree in industrial engineering from the University of California at Berkeley, and the Ph.D. degree in industrial engineering from Georgia Institute of Technology, in 2001. He is currently a Professor with the Department of Industrial Engineering. He is the Former Director of the Center on Stochastic Modeling, Optimization, & Statistics (COSMOS) and a Founder and the Director of the Healthcare Enterprise Research Center (HERC). His research interests include mathematical optimization and simulation in transportation, defense, and health care.

APURVA PAMIDIMUKKALA received the B.E. degree in civil engineering from Osmania University, Hyderabad, India, in 2018, and the Ph.D. degree in civil engineering from The University of Texas at Arlington, Arlington, TX, USA, in 2023. She is currently an Assistant Professor of research with The University of Texas at Arlington. Her research interests include construction safety, resilient infrastructure, and smart and sustainable mobility.