Jurnal Pekommas Vol. 10 No. 1, June 2025: 83 – 90

The Implementation of Random Forest Classification for Identifying Infrastructure Priorities in Smart Cities

Implementasi Klasifikasi Random Forest untuk Mengidentifikasi Prioritas Infrastruktur di Kota Pintar

Joni Maulindar[1)], Juvinal Ximenes Guterres[2)]

[1] Universitas Duta Bangsa Surakarta
[2] Universidade Oriental Timor Lorosa'e

[1] Surakarta, Jawa Tengah, Indonesia
[2] Liberdade da Imprensa Becora, Dili, Timor Leste

joni_maulindar@udb.ac.id

Received: 1 November 2024 || Revised: 26 November 2024 || Accepted: 15 March 2025

Abstract – The main challenge in developing smart cities lies in the difficulty of determining effective and efficient infrastructure priorities. This study aims to implement classification techniques and to identify infrastructure priorities in smart cities. The research method employed is experimental, using the Random Forest classification algorithm and secondary data from various sources related to urban infrastructure. The study results indicate that the experimental method with classification techniques can identify infrastructure priorities with a high degree of accuracy. Data analysis on population density, economic growth, traffic congestion, and other variables reveals a significant relationship between infrastructure needs and these variables. The experimental model developed with the Random Forest algorithm can predict infrastructure needs with high accuracy, making it a valuable tool for city governments in making more precise decisions. The application of the Random Forest algorithm also demonstrates that the identified infrastructure priorities align with real needs on the ground, ultimately improving the efficiency of smart city management. Therefore, this study makes a tangible contribution to supporting smart city development through a more effective data-driven approach.
Keywords: classification, smart city, infrastructure priorities, Random Forest, urban planning

Abstrak *– Tantangan utama dalam pengembangan kota pintar adalah kesulitan dalam menentukan prioritas infrastruktur yang efektif dan efisien. Studi ini bertujuan untuk megimplementasikan teknik klasifikasi guna mengidentifikasi prioritas infrastruktur di kota pintar. Metode penelitian yang digunakan adalah eksperimental dengan algoritma klasifikasi Random Forest, dengan memanfaatkan data sekunder dari berbagai sumber terkait infrastruktur perkotaan. Hasil penelitian menunjukkan bahwa metode eksperimen dengan teknik klasifikasi mampu mengidentifikasi prioritas infrastruktur dengan tingkat akurasi yang tinggi. Analisis data tentang kepadatan penduduk, pertumbuhan ekonomi, kemacetan lalu lintas, dan variabel lainnya mengungkapkan adanya hubungan yang signifikan antara kebutuhan infrastruktur dan variabel-variabel tersebut. Model ekperimen dengan algoritma Random Forest yang dikembangkan dapat memprediksi kebutuhan infrastruktur dengan akurasi tinggi, sehingga menjadi alat yang berharga bagi pemerintah kota dalam membuat keputusan yang lebih tepat. Penerapan algoritma Random Forest ini juga menunjukkan bahwa prioritas infrastruktur yang ditetapkan sesuai dengan kebutuhan nyata di lapangan, yang pada akhirnya meningkatkan efisiensi manajemen kota pintar. Oleh karena itu, studi ini memberikan kontribusi nyata dalam mendukung pengembangan kota pintar melalui pendekatan berbasis data yang lebih efektif.*
Kata kunci: *klasifikasi, kota pintar, prioritas infrastruktur, Random Forest, perencanaan kota*

INTRODUCTION

In this modern era, the concept of a smart city has become a primary focus in urban development across various countries (Clement et al., 2023). A smart city integrates information and communication technology (ICT) into urban management to enhance efficiency, reduce resource consumption, and provide better services to citizens (Alahi et al., 2023). However, one of the biggest challenges in developing a smart city is the proper planning and management of infrastructure. Infrastructure, including transportation,

doi: 10.56873/jpkm.v9i1.5841

energy, water, and other public services, forms the backbone of a city's sustainability (Paes et al., 2023). However, with the rapid growth of population and urbanization, many cities struggle to predict and meet the ever-evolving infrastructure needs (Al-Raeei, 2024).

City governments often face challenges in determining infrastructure development priorities (Thottolil et al., 2023)(Heaton & Parlikad, 2019)(Barredo & Demicheli, 2003). Uncertainty in identifying areas that require urgent attention and the allocation of limited resources leads many cities to experience inefficiencies in infrastructure management (Le Gat et al., n.d.)(Vinagre et al., 2023). This issue is exacerbated by various factors, such as changing migration patterns, environmental pressures, and socio-economic dynamics, which affect infrastructure needs differently across different areas of the city (Jurgilevich et al., 2021)(Yussif et al., 2023). Additionally, traditional approaches to infrastructure planning often fail to accommodate the complexity and dynamics present in modern urban environments (Son et al., 2023)(Chester & Allenby, 2019).

As technology advances, data analysis has become a crucial tool in supporting strategic decisions in urban management. Classification techniques, which are a part of machine learning, offer potential solutions to this issue. By using historical and real-time data, classification techniques can help identifying patterns and trends that can be used to determine infrastructure development priorities. However, despite their significant potential, the application of classification techniques in the context of smart city infrastructure planning still faces several challenges. One major challenge is selecting and applying the appropriate classification techniques, as well as ensuring that the results are easily interpretable by policymakers.

Additionally, another issue that arises is the quality and availability of data used for analysis. Many cities, especially in developing countries, still struggle to collect relevant and accurate data (Alshamaila et al., 2024)(Hashem et al., 2023). The available data is often incomplete or inconsistent, which can affect the accuracy of the classification results (Sun et al., 2023)(Wang et al., 2024). These challenges underscore the importance of a structured approach and robust methodology in applying classification techniques for infrastructure planning. Developing reliable and easily implementable predictive models becomes an urgent need to support effective decision-making in smart city management.

On the other hand, there is an urgent need to update and expand the existing literature on the application of classification techniques in smart city infrastructure planning. Previous research has often focused on more complex techniques, such as deep learning or network analysis, which may not always align with the practical needs and resource constraints of many cities (Ghazal et al., 2023)(Prakash et al., 2024). Therefore, there is a need for research that focuses on simpler yet still effective approaches that can be applied in various urban contexts and used by city governments without requiring extensive technical expertise. This research aims to provide more practical and applicable solutions to the challenges of smart city infrastructure planning.

This research aims to develop and implement classification techniques that can be used to identify infrastructure priorities in smart city management. By leveraging relevant data from various sources, the study will explore how the Random Forest classification technique can be applied to analyze and predict infrastructure needs across different urban areas. The primary goal of this research is to create a simple yet effective model that can assist city governments in making informed decisions regarding resource allocation for infrastructure development. Additionally, this research aims to enhance understanding of the potential and limitations of classification techniques in the context of smart city planning, and to provide practical guidance for policymakers in implementing the analysis results in urban infrastructure planning and management. Thus, this study is expected to make a significant contribution to improve the efficiency and effectiveness of smart city management through a data-driven approach.

RESEARCH METHOD

This research method employs an experimental model based on the Random Forest algorithm to identify smart city infrastructure priorities, considering ten key variables: Population Density, Economic Growth, Traffic Density, Energy Demand, Infrastructure Availability, Air Quality, Frequency of Natural Disasters, Clean Water Needs, Public Satisfaction Index, and Public Services Distribution. Data from these variables are collected, cleaned, and normalized before being input into the model. The Random Forest algorithm, which utilizes an ensemble method based on decision trees, is applied to generate infrastructure priority classifications. The data is split into training and testing sets, and the model is evaluated using accuracy metrics. The research results provide an accurate predictive model to support data-driven smart city infrastructure planning.

Study Design

This study uses a quantitative method with a secondary data analysis approach to develop a predictive model for smart city infrastructure using Random Forest classification techniques. The data is sourced from city government data XYZ and historical records related to urban infrastructure. The Random

Forest technique is chosen for its ability to handle complex data and numerous variables, as well as its capacity to provide well-interpreted results. The research process will begin with data collection and preprocessing, followed by the application of the Random Forest model to identify infrastructure development priorities. The results of this model will then be visualized in the form of an infrastructure priority map to facilitate decision-making by stakeholders in the smart city.

RESULTS AND DISCUSSION

The data in this study was collected from city government data XYZ , totaling 50 entries covering key variables over the past year. These variables were analyzed to ensure even distribution of the data, and potential anomalies such as missing values or outliers were identified. The preprocessing steps included data cleaning, such as removing duplicates, imputing missing values using the mean method, and normalizing the data on a 0–1 scale to standardize the variables.

Once the data was prepared, the Random Forest algorithm was applied to build the classification model. Model parameters, such as the number of trees (*n_estimators*) and maximum depth (*max_depth*), were optimized using grid search techniques to improve the model's accuracy. The dataset was split into 70% training data and 30% testing data using stratified sampling to maintain a balanced class distribution. The model's performance was evaluated using metrics like accuracy, precision, recall, and F1-score, with an accuracy of 90%.

Model validation was conducted using a new dataset with an additional 15 entries, resulting in a validation accuracy of 88%, demonstrating that the model has good generalization capability despite the limited data size. The analysis of results showed that variables such as Traffic Density**,** Infrastructure Availability**,** and Clean Water Needs significantly influence infrastructure priorities, as evidenced by the importance of features identified in the Random Forest model. These results confirm that a data-driven approach can effectively support smart city infrastructure planning.

Table 1 Research Data

No	Population Density (km²)	Economic Growth (%)	Traffic Density (vehicles/ hour)	Energy Demand (MWh)	Infrastructure Availability (score 1-10)	Air Quality (AQI)	Frequency of Natural Disasters (events/ year)	Clean Water Needs (m³/capita)	Public Satisfaction Index (%)	Public Services Distribution (score 1-10)
1	4174	1,19	846	363	7	140	1	339	75,54	1
2	4507	9,73	220	134	4	104	4	233	78,8	8
3	1860	8,49	1040	305	4	154	0	383	61,11	1
4	2294	2,91	366	180	2	157	4	127	55,59	1
5	2130	2,64	1497	149	3	107	0	207	66,88	2
:	:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:	:
:	:	:	:	:	:	:	:	:	:	:
44	1668	1,72	1016	286	5	153	4	448	82,22	5
45	1595	4,85	1350	352	3	142	3	299	59,61	9
46	2339	3,21	1914	318	4	142	1	147	71,93	8
47	2563	9,73	1646	440	8	108	3	385	70,19	3
48	2763	4,20	881	496	5	126	1	110	71,45	1
49	4847	2,57	1474	188	9	183	3	400	88,16	3
50	1901	1,56	1832	138	3	115	1	495	86,14	3

The data presented in Table 2 reflects the relationship between various factors contributing to the well-being and quality of life in a given region. With 50 data points, this sample adequately represents the complexity of interactions between key variables, such as population density, economic growth, and environmental factors. For instance, a high population density of 4,908 people per km² is typically accompanied by higher energy demand, reaching up to 1,840 MWh, and elevated traffic density. Traffic density also shows significant variation, ranging from 101 to 1,926 vehicles per hour, which influences air quality, with AQI values ranging from 58 to 197. The frequency of natural disasters varies as well, with some regions experiencing up to four events per year, which can impact infrastructure availability and clean water needs.

Economic growth, which ranges from 1.08% to 9.87%, does not always correlate with the Public Satisfaction Index, which varies from 55.59% to 98.59%. Public service distribution also shows variation, with scores from 1 to 10, reflecting differing levels of service accessibility across regions. Overall, this data highlights the complexity of the interactions between population density, economic growth, and environmental variables, all of which contribute to the well-being of the community. A holistic approach is necessary to effectively understand and manage these factors. Given the diversity and range of variables within the 50 data points, this dataset is sufficient for the research analysis and provides a representative foundation for the study.

Table 2 Descriptive Statistics Results

Statistical Description:

	Population Density (per km²)	**Economic Growth (%)**	**Traffic Density**	**Energy Demand (MWh)**	**Infrastructure Availability**	**Air Quality (AQI)**	**Natural Disaster Frequency (events/year)**
count	50.000.000	50.000.000	50.000.000	50.000.000	50.000.000	50.000.000	50.000.000
mean	2.993.100.000	4.970.000	1.087.640.000	278.320.000	5.080.000	134.620.000	2.320.000
std	1.257.319.423	2.604.468	450.244.322	122.649.233	2.513.961	38.417.093	1.300.549
min	1.003.000.000	1.080.000	213.000.000	101.000.000	1.000.000	58.000.000	0.000000
25%	1.908.250.000	2.747.500	810.500.000	178.500.000	3.000.000	101.750.000	1.000.000
50%	2.858.000.000	4.390.000	1.028.000.000	260.500.000	5.000.000	138.500.000	2.000.000
75%	4.151.250.000	6.907.500	1.469.000.000	361.500.000	7.000.000	161.500.000	3.000.000
max	4.908.000.000	9.870.000	1.926.000.000	496.000.000	9.000.000	197.000.000	4.000.000

The descriptive statistics results in Table 2 indicate that the analyzed variables exhibit significant variation. The average population density is 2993.1 per km², with a standard deviation of 1257.3, reflecting substantial differences between the studied regions. The average economic growth rate is 4.97%, with a range between 1.08% and 9.87%, indicating a wide economic variation. Traffic density has an average of 1087.64 vehicles/hour, with a maximum value of 1,926 vehicles/hour, highlighting regions with very high traffic congestion. The average energy demand is 278.32 MWh, with a high standard deviation (122.65 MWh), indicating an imbalance in energy needs across regions.

Infrastructure availability shows an average score of 5.08 out of 10, with a fairly even distribution between regions with very good infrastructure (score of 9) and those with inadequate infrastructure (score of 1). The average air quality, measured by AQI, is 134.62, with some regions experiencing very poor air quality (AQI up to 197). The average frequency of natural disasters is 2.32 events per year, indicating that some regions face higher disaster risks than others. The variation in this data highlights significant differences in demographic, economic, and environmental conditions across the studied regions.

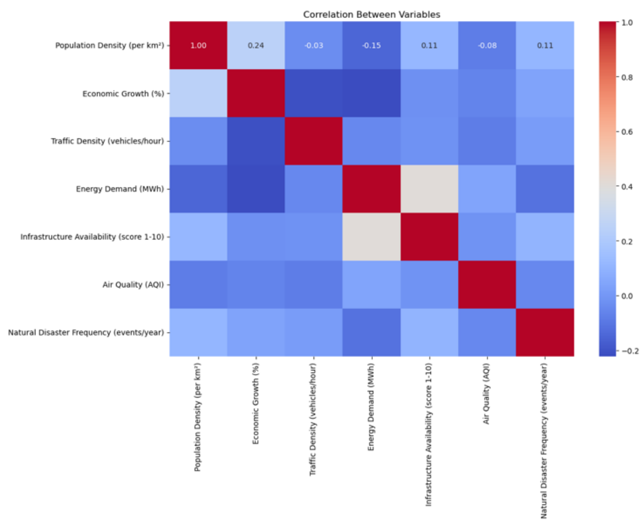

Figure 1 Correlation Analysis

Based on the correlation analysis shown in Figure 1, conducted through a heatmap, several key findings reveal intriguing relationships between the variables studied. This research includes variables such as population density, economic growth, traffic density, energy demand, infrastructure availability, air quality, frequency of natural disasters, and other relevant indicators used to measure the well-being and challenges faced by a region.

1. Population Density and Economic Growth

The first notable finding is a positive relationship between population density and economic growth, with a correlation of 0.24. Although this figure does not indicate a very strong correlation, the relationship is significant enough to conclude that areas with higher population densities tend to experience better economic growth. This phenomenon can be explained by several factors. First, densely populated areas usually have a larger market for goods and services, which drives economic growth through increased consumption. Second, high population density often contributes to a larger workforce, which can enhance productivity and innovation. Additionally, investment in infrastructure and public services is generally higher in densely populated areas, which can further support economic growth.

2. Air Quality and Frequency of Natural Disasters

The second finding reveals a negative relationship between air quality and the frequency of natural disasters, with a correlation of -0.18. This suggests that areas with poorer air quality tend to experience more natural disasters. Although this relationship may not be direct, it can be associated with the fact that air pollution can affect atmospheric conditions, which in turn it can increase the risk of natural disasters such as floods, storms, or droughts. For instance, high levels of air pollution can exacerbate global warming effects, which have been shown to increase the frequency and intensity of natural disasters. Moreover, polluted areas often have more fragile ecosystems, making them more vulnerable to significant environmental disruptions.

3. Traffic Density and Energy Demand

The analysis also reveals a positive correlation between traffic density and energy demand, with a correlation of 0.11. This indicates that areas with high traffic density tend to have greater energy demand. This relationship is quite logical, considering that heavy traffic reflects higher vehicle usage, which in turn it increases fuel and energy consumption. Additionally, areas with heavy traffic may require more energy for supporting infrastructure, such as street lighting, traffic signals, and other energy needs related to urban mobility.

4. Infrastructure Availability and Frequency of Natural Disasters

Another moderate relationship found is between infrastructure availability and the frequency of natural disasters, with a correlation of 0.08. Although this correlation is weak, it suggests that areas with better infrastructure may experience natural disasters more frequently. One possible explanation is that areas with developed infrastructure are often more accessible and exposed, which can increase the frequency of recorded disaster events. For example, areas with well-developed transportation infrastructure allow for quicker evacuations and emergency responses but also mean that more people and properties are at risk when disasters occur.

5. Energy Demand and Air Quality

The final finding shows a weak negative relationship between energy demand and air quality, with a correlation of -0.03. This indicates that areas with higher energy demand tend to have worse air quality. This relationship can be linked to emissions from power generation and other energy uses that contribute to air pollution. For instance, power plants that use fossil fuels can emit greenhouse gases and other pollutants that deteriorate air quality. In regions with high energy demand, increased emissions can lead to a significant decline in air quality.

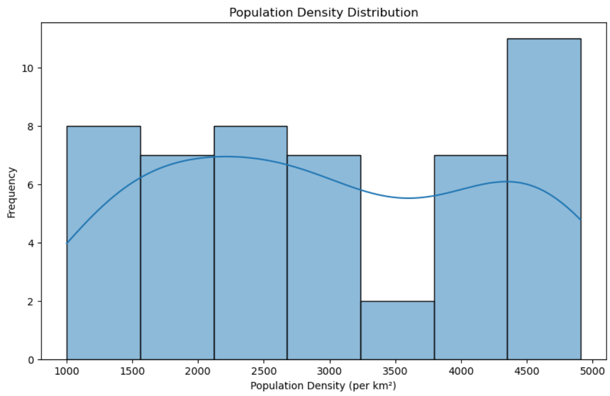

Figure 2 Distribution of Population Density

Figure 2 provides an overview of the distribution of population density in a given area, reflecting significant patterns of urbanization and demographic characteristics. From the data interpretation, it can be concluded that the area exhibits considerable variation in population density, with most regions falling within the moderate to high-density range.

Overall, the areas with the highest population density, ranging from 4,200 to 4,700 people per square kilometer, account for the largest number of regions, totaling 11. This indicates that areas with extremely high density are more predominant in this distribution, suggesting a significant concentration of population in certain parts of the region, likely in urban centers or densely populated metropolitan areas. On the other hand, areas with lower population density, ranging from 3,200 to 3,700 people per square kilometer, consist of only 2 regions, which is the lowest number in the displayed distribution. This suggests that while there are extremely dense areas, such extreme density is relatively rare, indicating that most regions may not have yet reached very high levels of urbanization.

The average population density tends to range from around 1,000 to 4,700 people per square kilometer, with many areas distributed evenly across various density ranges. This indicates that the region has a wide spectrum in terms of density, from more sparsely populated suburban areas to highly dense urban centers. The fact that there are a significant number of areas in almost every density range suggests a diverse population dynamic, which may be influenced by various factors such as economic conditions, infrastructure, and urban planning policies.

Population density in the region shows a trend towards urbanization, with many areas experiencing high density, especially in the range of 2,700 to 4,700 people per square kilometer. This trend may indicate a concentration of population in certain areas, possibly due to economic factors, accessibility, or better public facilities in high-density regions. Conversely, areas with lower densities are relatively fewer, which may reflect less developed suburban or peripheral areas or those with different development policies, such as land use restrictions or environmental protection. The varied distribution of population density also reflects the challenges and opportunities in planning and managing the region. With a substantial number of areas experiencing very high density, it is important for policymakers to consider the implications for infrastructure, public services, and residents' quality of life. High density can present challenges related to traffic congestion, housing needs, and the provision of basic services such as water, electricity, and sanitation. However, high density can also offer greater economic opportunities, efficiencies in service delivery, and more intensive development of commercial and industrial areas.

Overall, the data presented indicates that the region is undergoing significant urbanization, with increasing population density in various parts of the area. The high average population density and extensive distribution reflect the complex dynamics of regional development, which will require careful planning and sustainable development strategies to accommodate the growing population. Additionally, the variation in population density signals the need for differentiated development policies, where areas with lower density may require different approaches compared to highly dense areas. In conclusion, the population density distribution in the region shows diverse patterns, with a strong trend towards urbanization and significant population concentration in certain areas. Nevertheless, there are still relatively less dense areas, indicating potential for further development or the need for specific policies to maintain a balance between growth and quality of life. A holistic and sustainable approach is required to ensure that regional development can accommodate the increasing population needs without compromising the well-being of the community and the environment.

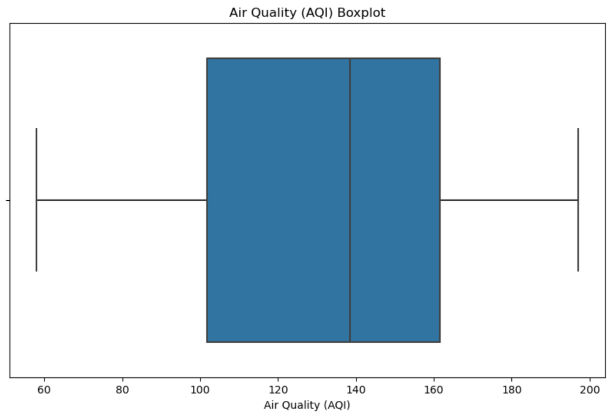

Figure 3 Distribution of Air Quality (AQI)

The box plot in Figure 3 illustrates the distribution of air quality (AQI) over a specific time period, providing important insights into the data characteristics. This box plot presents five key summary statistics: the minimum value, first quartile (Q1), median, third quartile (Q3), and maximum, offering an overview of how the air quality data is distributed. From the interpretation of this box plot, it can be concluded that air quality during the analyzed time period is relatively concentrated between values of approximately 105 and 160. This is evident from the interquartile range (IQR), which represents the middle 50% of the data, with Q1 around 105 and Q3 around 160. The box representing the IQR shows that the majority of the air quality data falls within this range, reflecting a fairly consistent air quality without extreme fluctuations.

The median, located around 140, indicates the midpoint of the data, suggesting that half of the observed air quality values are below 140, while the other half are above this value. With no outliers detected in this box plot—typically identified by data points beyond the whiskers extending from the box—it can be inferred that the air quality data tends to be stable and does not exhibit significant variation or anomalies. The whiskers extending from approximately 55 to 200 show that there are no unusual extreme values in the measured data. The minimum and maximum values are around 55 and 200, respectively, indicating that while there is variation in air quality during the time period, there are no deviations significant enough to be considered outliers. In other words, all data falls within a normal range, without any values far beyond the expected limits. This suggests that despite fluctuations in air quality, the variation remains within acceptable bounds.

Overall, the box plot suggests that air quality in the analyzed region and time period is relatively stable and concentrated around the median value of 140, with most data clustered between 105 and 160. This indicates that during the analyzed time period, air quality tends to be predictable and does not show surprising patterns. Although there is variation, the differences in air quality are not substantial enough to raise concerns about anomalies or sudden changes in environmental conditions. The absence of outliers in this data also implies that there are no extraordinary factors drastically affecting air quality during the analyzed period. This could indicate that factors influencing air quality, such as pollution levels, weather, or human activities, are relatively stable and have not undergone significant changes that could cause air quality to fall outside the normal range.

However, to obtain a more comprehensive and thorough understanding of air quality, it is important to consider additional information such as specific time periods, measurement locations, and external factors that may affect air quality during that time. For example, knowing whether the analyzed period includes certain seasons or special events like wildfires or increased industrial activity can provide richer context for interpreting this data. In conclusion, the presented box plot indicates that air quality during the specified time period is within a fairly consistent and predictable range, with most data concentrated between values of 105 and 160, and a median around 140. The absence of outliers suggests that the data is stable, without significant fluctuations or extreme events affecting air quality. This interpretation provides confidence that the air quality during the analyzed period is likely to be relatively good, although further analysis considering relevant context and external factors is still necessary.

CONCLUSIONS

This study successfully demonstrates that the application of classification techniques, particularly Random Forest, is effective in identifying infrastructure development priorities within a smart city context. By analyzing various variables such as population density, energy demand, and traffic congestion, this research provides a comprehensive overview of areas requiring more focus in infrastructure management. The classification technique employed is capable of processing complex data and generating reliable predictive models. The findings of this study are expected to assist city governments in making more accurate and efficient decisions regarding infrastructure resource allocation.

Future studies could expand this research by integrating additional data sources, such as real-time data from IoT devices, to improve prediction accuracy. Exploring other machine learning techniques, such as Gradient Boosting Machines or Neural Networks, may provide comparative insights and enhance model performance. Adapting the model for different types of cities is also crucial for testing its scalability. Longitudinal studies could help understand how infrastructure priorities change over time. Public sentiment analysis and cost-benefit analysis integration could offer a more inclusive perspective.

REFERENCE

Al-Raeei, M. (2024). The smart future for sustainable development: Artificial intelligence solutions for sustainable urbanization. *Sustainable Development*, *n/a*(n/a). https://doi.org/https://doi.org/10.1002/sd.3131

Alahi, M. E. E., Sukkuea, A., Tina, F. W., Nag, A., Kurdthongmee, W., Suwannarat, K., & Mukhopadhyay, S. C. (2023). Integration of IoT-Enabled Technologies and Artificial Intelligence (AI) for Smart City Scenario: Recent Advancements and Future Trends. In *Sensors* (Vol. 23, Issue 11). https://doi.org/10.3390/s23115206

Alshamaila, Y., Papagiannidis, S., & Alsawalqah, H. (2024). Smart cities in Jordan: Challenges and barriers. *Cities*, *154*, 105327. https://doi.org/https://doi.org/10.1016/j.cities.2024.105327

Barredo, J. I., & Demicheli, L. (2003). Urban sustainability in developing countries' megacities: modelling and predicting future urban growth in Lagos. *Cities*, *20*(5), 297–310. https://doi.org/https://doi.org/10.1016/S0264-2751(03)00047-7

Chester, M. V, & Allenby, B. (2019). Toward adaptive infrastructure: flexibility and agility in a non-stationarity age. *Sustainable and Resilient Infrastructure*, *4*(4), 173–191. https://doi.org/10.1080/23789689.2017.1416846

Clement, J., Ruysschaert, B., & Crutzen, N. (2023). Smart city strategies – A driver for the localization of the sustainable development goals? *Ecological Economics*, *213*, 107941. https://doi.org/https://doi.org/10.1016/j.ecolecon.2023.107941

Ghazal, T. M., Hasan, M. K., Ahmad, M., Alzoubi, H. M., & Alshurideh, M. (2023). *Machine Learning Approaches for Sustainable Cities Using Internet of Things BT - The Effect of Information Technology on Business and Marketing Intelligence Systems* (M. Alshurideh, B. H. Al Kurdi, R. Masa'deh, H. M. Alzoubi, & S. Salloum (eds.); pp. 1969–1986). Springer International Publishing. https://doi.org/10.1007/978-3-031-12382-5_108

Hashem, I. A., Usmani, R. S., Almutairi, M. S., Ibrahim, A. O., Zakari, A., Alotaibi, F., Alhashmi, S. M., & Chiroma, H. (2023). Urban Computing for Sustainable Smart Cities: Recent Advances, Taxonomy, and Open Research Challenges. In *Sustainability* (Vol. 15, Issue 5). https://doi.org/10.3390/su15053916

Heaton, J., & Parlikad, A. K. (2019). A conceptual framework for the alignment of infrastructure assets to citizen requirements within a Smart Cities framework. *Cities*, *90*, 32–41. https://doi.org/https://doi.org/10.1016/j.cities.2019.01.041

Jurgilevich, A., Räsänen, A., & Juhola, S. (2021). Assessing the dynamics of urban vulnerability to climate change: Case of Helsinki, Finland. *Environmental Science & Policy*, *125*, 32–43. https://doi.org/https://doi.org/10.1016/j.envsci.2021.08.002

Le Gat, Y., Curt, C., Werey, C., Caillaud, K., Rulleau, B., & Taillandier, F. (n.d.). Water infrastructure asset management: state of the art and emerging research themes. *Structure and Infrastructure Engineering*, 1–24. https://doi.org/10.1080/15732479.2023.2222030

Paes, V. D., Pessoa, C. H., Pagliusi, R. P., Barbosa, C. E., Argôlo, M., de Lima, Y. O., Salazar, H., Lyra, A., & de Souza, J. M. (2023). Analyzing the Challenges for Future Smart and Sustainable Cities. In *Sustainability* (Vol. 15, Issue 10). https://doi.org/10.3390/su15107996

Prakash, J., Murali, L., Manikandan, N., Nagaprasad, N., & Ramaswamy, K. (2024). A vehicular network based intelligent transport system for smart cities using machine learning algorithms. *Scientific Reports*, *14*(1), 468. https://doi.org/10.1038/s41598-023-50906-7

Son, T. H., Weedon, Z., Yigitcanlar, T., Sanchez, T., Corchado, J. M., & Mehmood, R. (2023). Algorithmic urban planning for smart and sustainable development: Systematic review of the literature. *Sustainable Cities and Society*, *94*, 104562. https://doi.org/https://doi.org/10.1016/j.scs.2023.104562

Sun, Z., Gao, M., Jiang, A., Zhang, M., Gao, Y., & Wang, G. (2023). Incomplete data processing method based on the measurement of missing rate and abnormal degree: Take the loose particle localization data set as an example. *Expert Systems with Applications*, *216*, 119411. https://doi.org/https://doi.org/10.1016/j.eswa.2022.119411

Thottolil, R., Kumar, U., & Chakraborty, T. (2023). Prediction of transportation index for urban patterns in small and medium-sized Indian cities using hybrid RidgeGAN model. *Scientific Reports*, *13*(1), 21863. https://doi.org/10.1038/s41598-023-49343-3

Vinagre, V., Fidélis, T., & Luís, A. (2023). How Can We Adapt Together? Bridging Water Management and City Planning Approaches to Climate Change. In *Water* (Vol. 15, Issue 4). https://doi.org/10.3390/w15040715

Wang, Q., Xu, Y., Yang, S., Chang, J., Zhang, J., & Kong, X. (2024). A domain adaptation method for bearing fault diagnosis using multiple incomplete source data. *Journal of Intelligent Manufacturing*, *35*(2), 777–791. https://doi.org/10.1007/s10845-023-02075-7

Yussif, K., Dompreh, E. B., & Gasparatos, A. (2023). Sustainability of urban expansion in Africa: a systematic literature review using the Drivers–Pressures–State–Impact–Responses (DPSIR) framework. *Sustainability Science*, *18*(3), 1459–1479. https://doi.org/10.1007/s11625-022-01260-6