SvedbergOpen
DISSEMINATION OF KNOWLEDGE

International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: https://www.svedbergopen.com/

Research Paper Open

AI-Based Decision Support System for Sustainable Infrastructure Development

1*Ashish Kumar Mishra, 2Mrigana Walia, 3Dr. Poorva Sanjay Sabnis, 4Mrs. Najiya Koderi Valappil, 5Dr. Megha Bansal, 6Abir.Mahmud Dipto

1*Independent Scholar, Department of Computer Science. E-mail: akmishra.cse13@nituk.ac.in, orcid: https://orcid.org/0009-0000-0437-549X
2Assistant Professor, CSE Department, Quest Infosys Foundation Group of Institutions, Zirakpur, Punjab, India, E-mail: mrigana9@gmail.com
3Assistant Professor, Computer Science and Engineering, St. Vincent Pallotti College of Engineering and Technology, Nagpur, India. E-mail: poorva.sabnis89@gmail.com, orcid: https://orcid.org/0000-0002-0490-9452
4PhD Scholar, Department of Electronics, Banasthali Vidhyapith, Rajasthan, India.
E-mail: najiyakoderi@gmail.com, orcid: https://orcid.org/0000-0001-7558-5718
5Assistant Professor, School of Information Technology, Vivekananda Institute of Professional Studies-Technical Campus, Delhi, India. E-mail: meghabansal05@gmail.com, orcid: https://orcid.org/0000-0002-1858-1726
6Researcher, Geotechnical Engineering, BUET-JIDPUS. E-mail: diptoabirruet.ce16@gmail.com, orcid: https://orcid.org/0009-0002-7499-3781

Abstract
Sustainable infrastructure development requires integrated assessment of economic, environmental, energy, and social conditions to support evidence-based planning and resource allocation. Existing studies have applied Machine Learning (ML) and Multi-Criteria Decision-Making (MCDM) techniques for sustainability assessment, but limited efforts integrate predictive modelling, explainability, objective weighting, and infrastructure ranking within a unified framework. This study develops an explainable Artificial Intelligence (AI)-based decision-support framework using global development indicators to identify influential factors and assess sustainable infrastructure performance across countries and entities. The framework comprises data preprocessing and screening, Recursive Feature Elimination (RFE), Light Gradient Boosting Machine (LightGBM)-based prediction, SHapley Additive exPlanations (SHAP)-based interpretation, Logarithmic Percentage Change-driven Objective Weighting (LOPCOW), and Evaluation based on Removal Effects of Criteria and Uncertainty-based Normalization and Scoring (ERUNS)-based ranking, followed by Technique for Order Preference by Similarity to Ideal Solution (TOPSIS) validation and sensitivity analysis. RFE identified 10 priority indicators, achieving a validation Root Mean Square Error (RMSE) of 2.34. LightGBM achieved R^2 = 0.9925. SHAP identified urban electricity access (18.783) and rural electricity access (5.649) as the most influential features. The framework provides an interpretable basis for sustainable infrastructure assessment and supports transparent, data-driven planning and prioritization.

Keywords: Artificial intelligence, Decision support, Explainable AI, LightGBM, Sustainable infrastructure

1. Introduction

Sustainable infrastructure development has become an essential component of long-term economic growth, environmental protection, and social well-being. Infrastructure systems comprising the building blocks of power, transport, energy, information, water, and sanitation systems shape and enhance the quality of life and the productive potential of nations [1-3]. The development of all these necessary infrastructures is constrained by urbanization, increasing energy demands, environmental degradation, and limited resources and fair distribution of basic services. The above challenges create a need for a systematic approach to analyzing infrastructure through a combination of several indicators of economy, environment, energy, and society [4-6]. Many assessment methods in use rely on single indicators or aggregated statistical measurements which capture sustainability dimensions only in a summary way. For this reason, flexibility and the potential to adapt to different conditions of analysis within data-centric models become increasingly important for developing decision-support frameworks that prioritize infrastructure and assess the impacts of choices [7,8]. The infrastructure life cycle, as shown in Figure 1, includes upstream activities such as strategic planning, priority setting, design, and project planning, as well as downstream financing, implementation, services, and infrastructure maintenance, and the closure or adaptive reuse of an infrastructure facility.

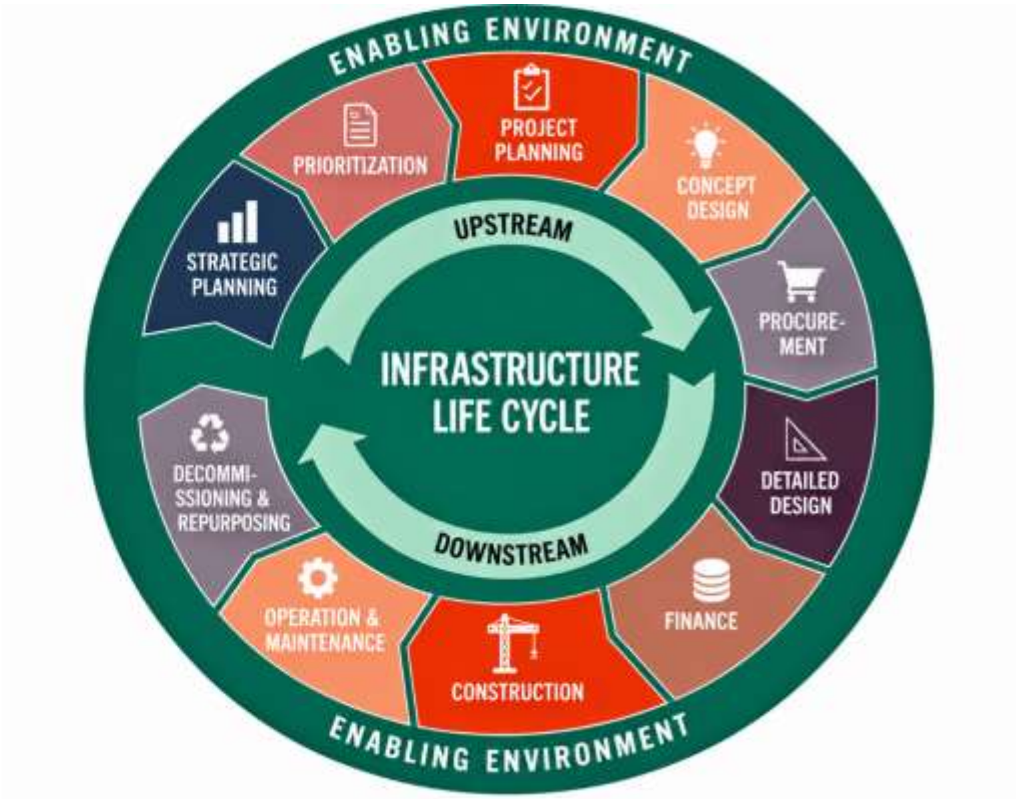

Figure 1: Sustainable infrastructure life cycle [9]

The proliferation of global development datasets and advancements in AI and ML show potential to strengthen sustainable infrastructure assessment. ML can identify complex relationships, such as nonlinear relationships, in socioeconomic, environmental, and infrastructure data, and develop robust predictive models for a range of development challenges [10-13]. Predictive accuracy, however, is not sufficient to describe the reasons behind the decision of a particular indicator in a particular situation (or vice versa), thereby constraining the use of ML in the framing of policy and planning for infrastructure [14]. Interpretability can be improved with the application of SHAP values in Explainable AI (XAI) to derive insights into the contribution of individual indicators to the prediction of the model. Simultaneously, MCDM methods can incorporate diverse dimensions of sustainability to create incremental rankings of countries or entities [15-18]. Thus, considering the available tools, a comprehensive framework of sustainable infrastructure development could be built.

Previous studies have applied ML, XAI, and MCDM methods separately, or have combined the methods in limited ways, to assess sustainability, plan cities, and to make infrastructure-related decisions. However, there continues to be an important gap in the methodology to integrate feature selection, predictive modeling, explanation of models, weighting of criteria, and sustainability-focused ranking with indicators of development that can be compared on a global scale [19-22]. In addition, 'big data' tends to have a high number of attributes with a great deal of missing values and redundant indicators. This tends to negatively affect the reliability of model construction and criterion selection [23]. This study aims to provide a decision-support framework to design sustainable infrastructure using the principles of XAI and using Global Development Indicators. RFE is integrated into the framework to narrow down competing indicators, while other components of the framework are data preprocessing and screening, LightGBM, SHAP, LOPCOW, and ERUNS. This framework is able to identify the top indicators of sustainable infrastructure in a given context, assess their predictive importance, assign objective weights, and provide a relative comparison of countries/entities. The main contributions of this study are:

- Develops an integrated decision-support framework combining RFE, LightGBM, SHAP, LOPCOW, and ERUNS for sustainable infrastructure assessment.
- Identifies a compact set of high-priority infrastructure-related indicators from a large pool of global development indicators while addressing missingness and feature redundancy.
- Employs LightGBM and SHAP to provide both predictive modelling and transparent interpretation of the relative contribution and direction of selected indicators.

- Applies LOPCOW-based objective weighting and ERUNS-based ranking to compare sustainable infrastructure performance across entities, with ranking robustness assessed through TOPSIS validation and criterion-weight sensitivity analysis.

The remainder of this paper is organized into four main sections. Section 2 covers the literature review, the third section describes the methodology, section 4 presents the results and discussion, and Section 5 concludes the study.

2. Literature Review

The reviewed studies demonstrate the growing application of AI, ML, digital twins, and Decision-Support Systems (DSS) for sustainable infrastructure and urban development. **Anderson et al. (2026) [24]** proposed a decision-centric framework integrating AI, digital twins, sensing, data analytics, and operational decision-making for intelligent and sustainable infrastructure. The framework was validated through a tunnel engineering case study and further generalized to transportation, energy, and flood management, demonstrating improvements in decision consistency, responsiveness, and operational efficiency. Similarly, **Shohan et al. (2026) [25]** developed a machine-learning framework for construction sustainability prediction using 150 survey responses and 19 attributes. Support Vector Machine (SVM), RF, and XGB models were evaluated using nested 10×5-fold cross-validation, with RF achieving the best performance with a test R^2 of 0.734 and 78% classification accuracy. Waste generation, energy consumption, and project duration were identified as key predictors.

Building on the integration of AI with decision-support methodologies, **Wang and Ren (2025) [26]** developed a sustainable urban planning DSS combining RF-RFE, LOPCOW, ERUNS, objective weighting, and uncertainty-based ranking. The framework consistently identified BS3 as the optimal alternative among five urban-planning scenarios, demonstrating its potential for objective and robust decision-making. **Fedorka et al. (2025) [27]** further examined adaptive AI learning models for DSS applications in smart regions by integrating ML, Deep Learning (DL), and Reinforcement Learning (RL) across urban, energy, and security management. Their implementation reported 25% faster decision-making, 15–18% cost savings, and up to 30% improvement in prediction accuracy.

In the context of climate-resilient infrastructure, **Ghafari and Samaei (2025) [28]** proposed an integrated AI–Digital Twin framework for Palm Jumeirah that combined Building Information Modeling (BIM)/ Geographic Information System (GIS), Internet of Things (IoT), remote sensing, Finite Element Modelling (FEM), Random Forest (RF), and Long-Short Term Memory (LSTM). The framework achieved an R^2 of 0.91 and further quantified climate-related structural responses, including 28% higher seawall shear, 22 mm thermal expansion, and a failure probability exceeding 0.63 under the RCP 8.5 scenario. **Lutfiani et al. (2024) [29]** investigated AI-based approaches for improving urban infrastructure resource efficiency using six months of energy, transportation, water, and environmental data. Neural-network and deep-learning approaches improved energy efficiency from 70% to 85%, reduced travel time by 15%, and enhanced water-management efficiency by 15%.

Similarly, **Shulajkovska et al. (2024) [30]** developed the open-source Urbanite framework for sustainable urban mobility and decision-making by integrating Multi-Agent Transport Simulation (MATSim) microscopic traffic simulation, Decision Expert (DEXi) decision support, Orange-based ML, and advanced visualization. Evaluation across four European cities demonstrated that multi-label ML reduced policy-evaluation time by three orders of magnitude, while the semantic prediction approach achieved 91% accuracy. **Hadiyana and Ji-hoon (2024) [31]** examined AI-driven urban planning through global case studies incorporating data integration, ML, predictive modelling, and AI-based decision support. Their findings indicated reductions of 33.3% in travel time and 15% in energy consumption, together with a 23.5% improvement in air quality. Finally, **Tomar (2024) [32]** proposed a conceptual AI-based DSS for sustainable development policy and planning, integrating data management, predictive analytics, optimization, visualization, and feedback mechanisms to support resource allocation, risk forecasting, policy coordination, and adaptive governance.

While many advances have been made in the field of AI-based sustainable infrastructure, some research gaps still exist. Many research studies focus on certain infrastructure systems or case studies [24,28,30]. However, a focus is lacking on internationally comparable, multi-faceted development indicators. Additionally, the integration of feature selection, modeling, explanation of AI, objective weighting, robustness of sustainability ranking, and a unified DSS is still limited [25,27,32]. These gaps show the need for a unified framework for ranking countries and regions based on sustainable development indicators. This framework should allow users to analyze the indicators and prioritize sustainable infrastructure development.

3. Research Methodology

This study describes an AI-based decision support framework, XAI, to aid in the use of publicly available Sustainable Development Goals (SDG) indicators to facilitate the development of sustainable infrastructure. The framework integrates data preprocessing, feature selection, LightGBM, SHAP, and MCDM to aid the identification and evaluation of the performance of sustainable infrastructure. Figure 2 shows the proposed methodology flowchart of this study.

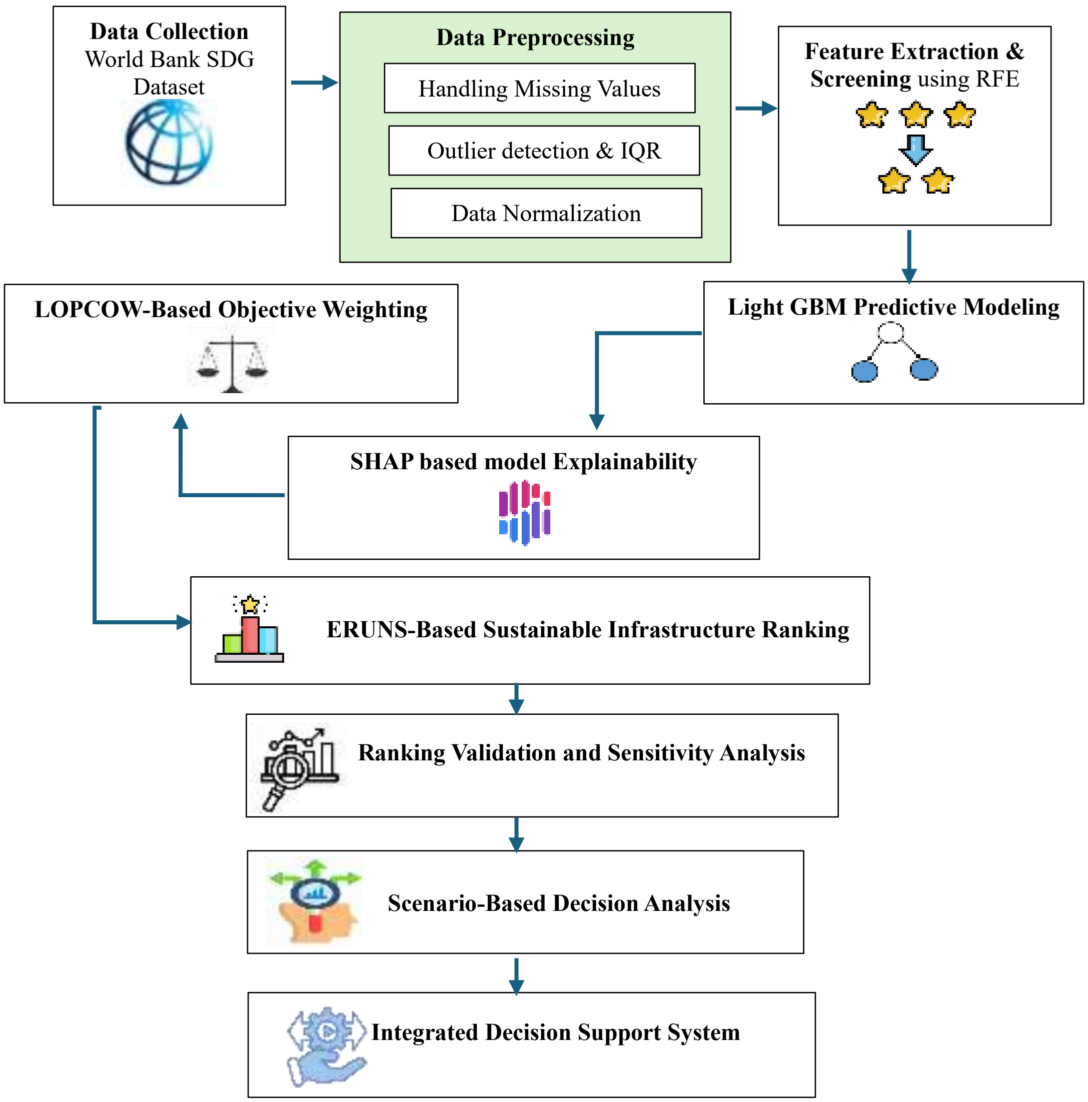

Figure 2: Proposed Methodology

3.1 Data Collection

This study uses the World Bank SDG dataset made available through Kaggle [33]. The dataset captures country-level development indicators across multiple sustainability dimensions. The dataset allows us to identify the specific infrastructure, environmental, economic, and social indicators for the DSS under development. Table 1 outlines the key attributes of the Dataset.

Table 1: Dataset Details

Attribute	Value
Dataset	SDG
Source	World Bank
Access	Kaggle
Data type	Secondary
Geographic level	Country

Coverage	Global
Period	1990–2020
Main domains	Economic, social, environmental, energy, health, ICT and urban development
Observation unit	Country-year
License	CC BY 4.0

The dataset is represented as a country–indicator matrix $X = [x_{ij}]$, where x_{ij} denotes the value of indicator j for country-year observation i.

3.2 Data Preprocessing

The purpose of preprocessing is to improve the quality of data with respect to completeness, consistency, and comparability prior to feature extraction and modeling. Three operations are applied: missing value treatment, outlier detection, and normalization. For each indicator, missing observations are analyzed. Indicators with a lot of missing values are removed, while indicators with a small number of missing values are filled using an appropriate method. An appropriate method is used to preserve as many observations as possible. Extreme outliers are identified using the Interquartile Range (IQR):

$$IQR = Q_3 - Q_1 \tag{1}$$

Where Q_1 and Q_3 are the first and third quartiles, respectively. Observations outside $Q_1 - 1.5IQR$ and $Q_3 + 1.5IQR$ are examined before treatment. The retained indicators are normalized to a common scale for MCDM analysis. Benefit and cost criteria are transformed using min–max normalization:

$$r_{ij} = \frac{x_{ij} - x_j^{min}}{x_j^{max} - x_j^{min}} \tag{2}$$

for benefit criteria, and

$$r_{ij} = \frac{x_j^{max} - x_{ij}}{x_j^{max} - x_j^{min}} \tag{3}$$

for cost criteria, where r_{ij} is the normalized value of indicator j for observation i.

3.3 Feature Extraction and Screening

Prior to analysis, the SDG indicators fall into sections of sustainable infrastructure, such as accessibility to infrastructure, energy, water and sanitation, digital connectivity, environmental and economic and social sustainability. Redundant indicators are found through Pearson correlation and stated as:

$$r_{jk} = \frac{\sum_{i=1}^{n}(x_{ij} - \bar{x}_j)(x_{ik} - \bar{x}_k)}{\sqrt{\sum_{i=1}^{n}(x_{ij} - \bar{x}_j)^2 \sum_{i=1}^{n}(x_{ik} - \bar{x}_k)^2}} \tag{4}$$

Where r_{jk} denotes the correlation between indicators j and k, x_{ij} represents the value of indicator j for observation i, and $\bar{x}_j$ denotes its mean. Indicators exhibiting a very high absolute correlation ($|r_{jk}|$) are screened to reduce redundancy, and the resulting candidate features are forwarded to the RFE-based feature selection stage.

3.4 RFE-Based Feature Selection

RFE is applied to choose the most important indicators from the first selected feature set. The LightGBM estimator removes the least important feature at each stage until the most suitable feature subset is obtained. The importance of features can be written as:

$$I_j = \sum_{t=1}^{T} G_{j,t} \tag{5}$$

Where I_j denotes the importance of feature j, T is the total number of LightGBM trees, and $G_{j,t}$ represents the gain contributed by feature j in tree t. Features with lower importance are recursively eliminated, while the subset producing the best validation performance is retained for predictive modelling.

3.5 LightGBM-Based Predictive Modelling

Features determined through RFE are used in the construction of the LightGBM model for predicting the sustainability infrastructure performance. LightGBM is a framework for creating an ensemble of gradient boosting decision trees. Each tree reduces the prediction error of the model as a whole. The prediction function is given by:

$$\hat{y}_i = \sum_{t=1}^{T} f_t(x_i),\ f_t \in F \tag{6}$$

Where $\hat{y}_i$ is the predicted outcome for observation i, T is the number of trees, f_t represents the t^{th} decision tree, x_i denotes the selected feature vector, and F is the space of decision trees. The model is trained on the selected indicators and optimized using validation data to obtain reliable predictive performance.

3.6 SHAP-Based Model Explainability

SHAP is used to analyze how each indicator contributes to the LightGBM predictions. SHAP decomposition for a prediction f(x) is defined as:

$$f(x) = \phi_0 + \sum_{j=1}^{p} \phi_j \quad (7)$$

where $f(x)$ is the model prediction, ϕ_0 is the expected model output, $p$is the number of selected features, and $\phi_j$represents the contribution of feature j to the prediction. The resulting SHAP values identify the most influential sustainability and infrastructure indicators and indicate whether their contributions increase or decrease the predicted performance, thereby improving the interpretability of the proposed DSS.

3.7 LOPCOW-Based Objective Weighting

The Logarithmic Percentage Change-driven Objective Weighting (LOPCOW) method is used to objectively weight the selected indicators. This method focuses on the relative importance of the criteria based on changes in the normalized decision matrix, diminishing the reliance on subjective expert judgments. The weight of criterion j is calculated as:

$$w_j = \frac{D_j}{\sum_{j=1}^{m} D_j} \quad (8)$$

Where w_j denotes the objective weight of criterion j, D_j represents its information/dispersion measure, and $m$is the total number of criteria. The resulting weights are used in the subsequent ERUNS-based evaluation.

3.8 ERUNS-Based Sustainable Infrastructure Ranking

To assess the overall sustainable infrastructure performance of countries using the Evaluation based on Relative Utility and Nonlinear Standardization (ERUNS) method for the LOPCOW criterion weights and integrated decision matrix, the overall utility score is the weighted score of the performance of the criteria.

$$S_i = \sum_{j=1}^{m} w_j \, u_{ij} \quad (9)$$

where S_i denotes the overall performance score of country i, w_j is the weight of criterion j, and u_{ij} represents the relative utility of country i with respect to criterion j. Countries are subsequently ranked according to their obtained performance scores, with higher scores indicating better sustainable infrastructure performance.

3.9 Ranking Validation and Sensitivity Analysis

The ranking system based on ERUNS is assessed using the TOPSIS method and sensitivity analysis. Using the TOPSIS method, country rankings are generated by their relative closeness to the ideal and/or negative-ideal solutions. Sensitivity analysis studies the impact of changes in the weighting of criteria on the rankings. The TOPSIS closeness coefficient is:

$$C_i = \frac{D_i^-}{D_i^+ + D_i^-} \quad (10)$$

Where C_i denotes the relative closeness of country i, D_i^+ represents its distance from the positive ideal solution, and D_i^- represents its distance from the negative ideal solution. Consistency between ERUNS and TOPSIS rankings and the stability of rankings under weight perturbations are used to assess the robustness of the proposed decision-support framework.

3.10 Integrated Decision Support System

To build an AI-based Decision Support System (DSS), outputs of the LightGBM predictive model, SHAP-based explanations, LOPCOW weights, and ERUNS ranking are integrated. The DSS integrates predictive and multi-criteria analyses to conduct a holistic appraisal of country-scale sustainable infrastructure development. For every country, DSS generates an overall performance score, performance rank relative to other countries, performance category, and Key Performance Indicators (KPIs) that strongly influence predicted performance. This integrated system helps to identify the country's infrastructure strengths and weaknesses and areas that require development investments.

3.11 Scenario-Based Decision Analysis

Scenario analysis is used to see how development priorities affect the ranking of sustainable infrastructures. Three development priority scenarios are designed by changing the weights of the criteria while holding the decision matrix constant. Business as usual (baseline) and ERUNS scores, as well as country rankings achieved under each priority, are examined along with the baseline to find out potential priorities for country ranking. It shows that the DSS developed for sustainable infrastructure planning is applicable for a variety of planning priorities.

3.12 Pseudocode

Algorithm Proposed_DSS(SDG_Dataset)

```
# Step 1: Data Collection
D ← Load_SDG_Dataset(SDG_Dataset)
D ← Select_Relevant_Countries_Years_Indicators(D)
# Step 2: Data Preprocessing
D ← Handle_Missing_Values(D)
D ← Detect_And_Treat_Outliers_IQR(D)
```

```
D ← Normalize_Indicators(D)
# Step 3: Feature Extraction and Screening
F ← Extract_Sustainable_Infrastructure_Features(D)
F ← Correlation_Screening(F)
# Step 4: RFE-Based Feature Selection
F_selected ← RFE(F, Estimator = LightGBM)
F_selected ← Select_Optimal_Feature_Subset(F_selected)
# Step 5: LightGBM Predictive Modelling
X_train, X_val, X_test, y_train, y_val, y_test ← Temporal_Split(D)
Model ← Train_LightGBM(X_train[F_selected], y_train)
Model ← Optimize_Model(Model, X_val[F_selected], y_val)
# Step 6: Model Evaluation
Performance ← Evaluate(Model, X_test[F_selected], y_test)
# Step 7: SHAP-Based Explainability
SHAP_Values ← Calculate_SHAP(Model, X_test[F_selected])
Global_Importance ← Determine_Global_Feature_Importance(SHAP_Values)
Local_Importance ← Determine_Country_Level_Contributions(SHAP_Values)
# Step 8: LOPCOW Objective Weighting
M ← Construct_Decision_Matrix(D, F_selected)
M ← Identify_Benefit_And_Cost_Criteria(M)
Weights ← LOPCOW(M)
# Step 9: ERUNS Ranking
Scores ← ERUNS(M, Weights)
Ranking ← Rank_Countries(Scores)
# Step 10: Ranking Validation and Sensitivity Analysis
TOPSIS_Ranking ← Apply_TOPSIS(M, Weights)
Ranking_Consistency ← Compare_Rankings(Ranking, TOPSIS_Ranking)
Sensitivity ← Perform_Sensitivity_Analysis(M, Weights)
# Step 11: Scenario Analysis
Environmental_Scenario ← Apply_Scenario(M, Environmental_Weights)
Social_Scenario ← Apply_Scenario(M, Social_Weights)
Economic_Scenario ← Apply_Scenario(M, Economic_Weights)
# Step 12: Integrated Decision Support
DSS_Output ← Integrate(
  Performance,
  Global_Importance,
  Local_Importance,
  Ranking,
  Ranking_Consistency,
  Sensitivity,
  Environmental_Scenario,
  Social_Scenario,
  Economic_Scenario
)
Return DSS_Output
```

4. Results Analysis

This section discusses and interprets the outcomes of data preprocessing, feature selection, LightGBM prediction, SHAP, LOPCOW weighting, and the ranking based on the ERUNS method. Ranking validation, sensitivity analysis, and scenario-based decision assessment are then used to analyze the robustness and the practical relevance of the framework. Table 2 shows the main features of the SDG dataset for the study, which was used for 263 countries/entities and 375 indicators over the years 1990 - 2018. The dataset contains 1,058,208 observed and 1,801,917 missing values. Overall, the missingness stands at 63.00%.

Table 2: Dataset Characteristics and Coverage

Attribute	Observed Value
Countries	263

Indicators	375
Study period available in uploaded file	1990–2018
Number of years	29
Possible observations	2,860,125
Available observations	1,058,208
Missing observations	1,801,917
Overall missingness	63.00%
Data structure	Country × Indicator × Year
Dataset type	Secondary, country-level

Figure 3 illustrates the distribution of indicators across different missingness levels. Most indicators fall within the 90–100% missingness range (117 indicators), while 56 and 55 indicators fall within the 10–20% and 80–90% ranges, respectively.

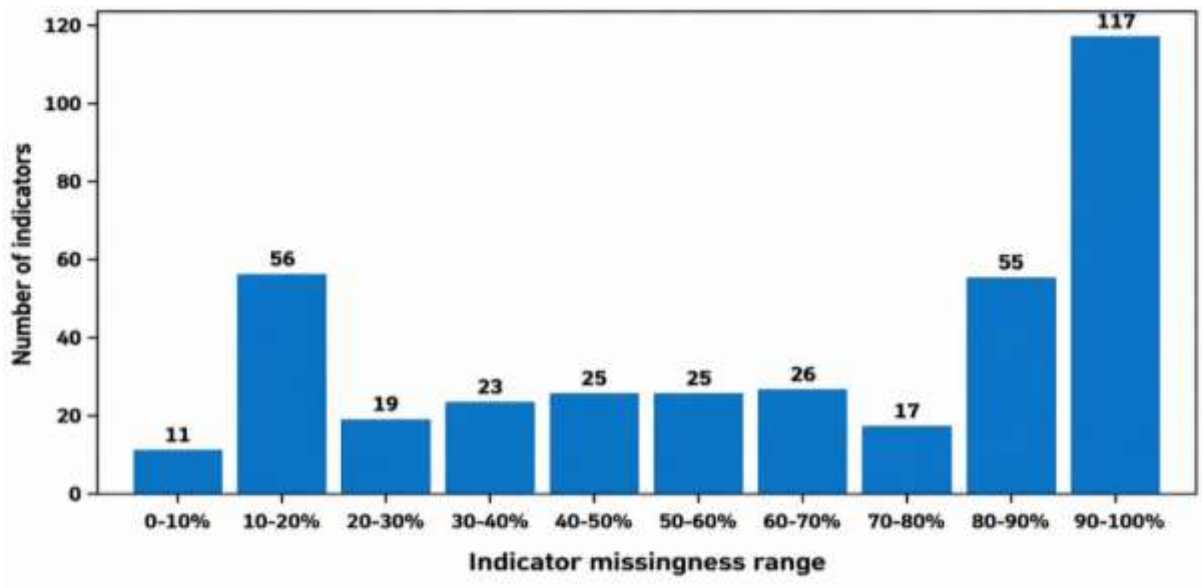

Figure 3: Distribution of Indicators by Missingness Level

Figure 4 illustrates the gap in data for each sustainable infrastructure indicator that was assessed. The indicators show a gap or 'missing' value for clean cooking access at 47.62%, while the gap for the urban population indicator is 4.26%.

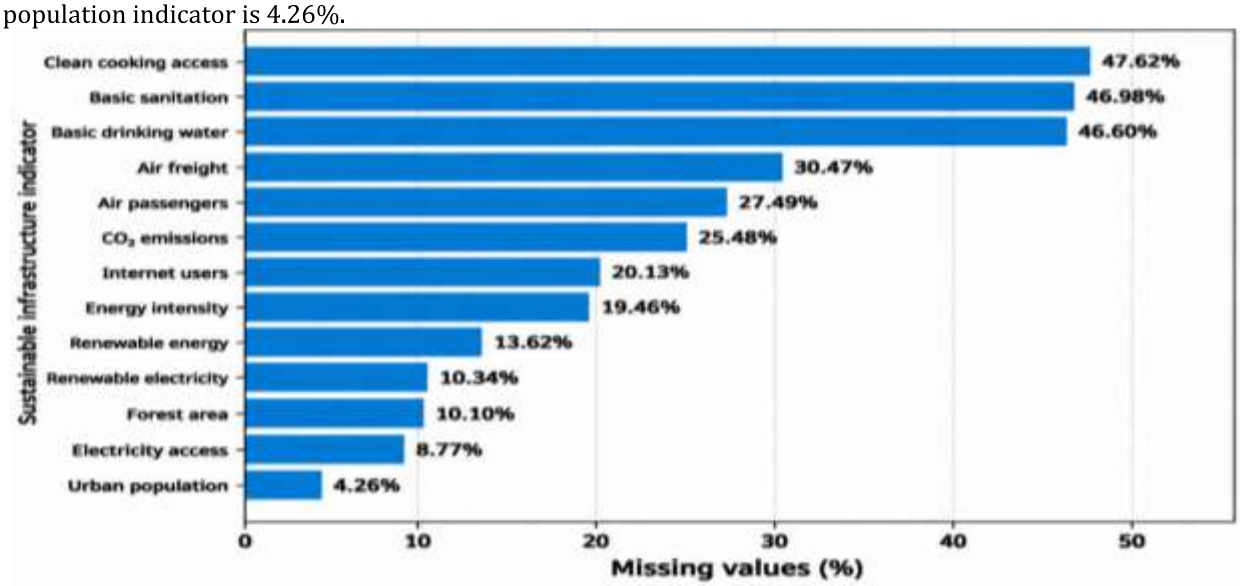

Figure 4: Indicator-Level Missingness after Screening

RMSE validation values for decreasing feature number, as determined by RFE, are shown in Figure 5. The value for RMSE decreased substantially from 12.67 for one feature to 2.34 for 10 features. Adding further features showed no significant benefit. Hence, choose 10 features as the optimal subset.

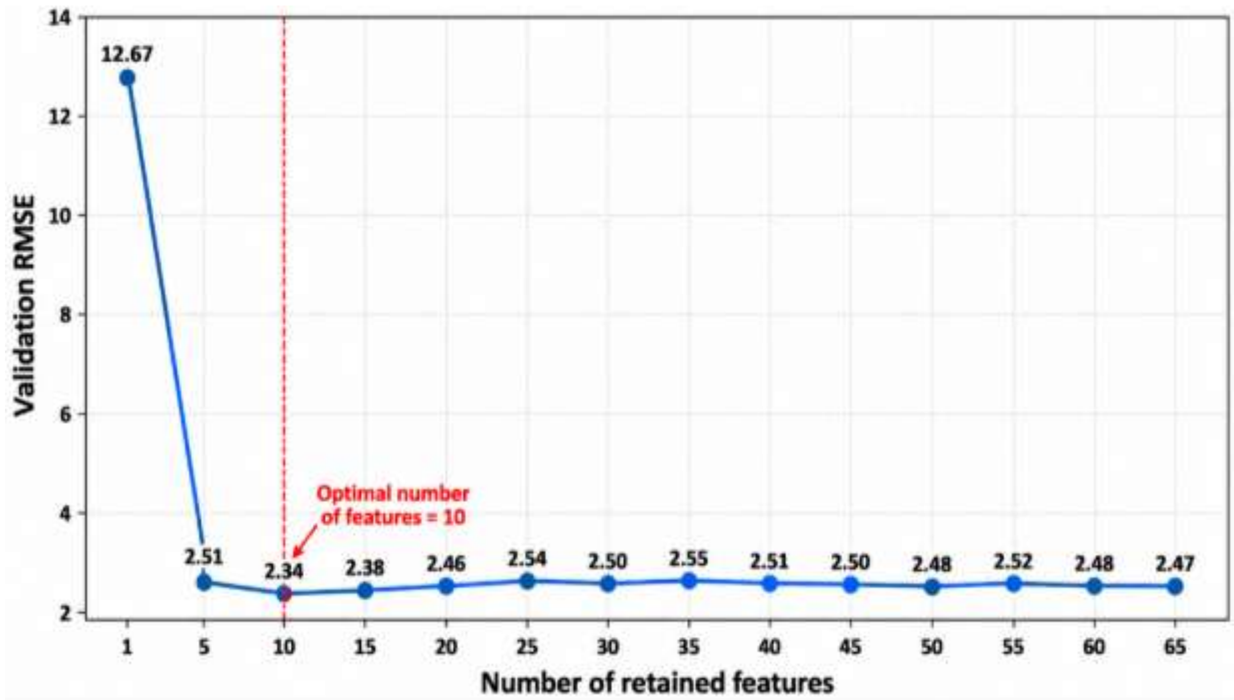

Figure 5: RFE Feature Selection Performance

Figure 6 ranks the top 20 indicators identified by RFE-rank 1: access to urban electricity. Access to electricity in rural areas was ranked 2. Other indicators of high priority were CO_2 emissions per capita (3), urban population (4), and urban population share (5). Other indicators of note were energy intensity (6) and renewable energy consumption (10), as well as internet use (15) and GDP (20).

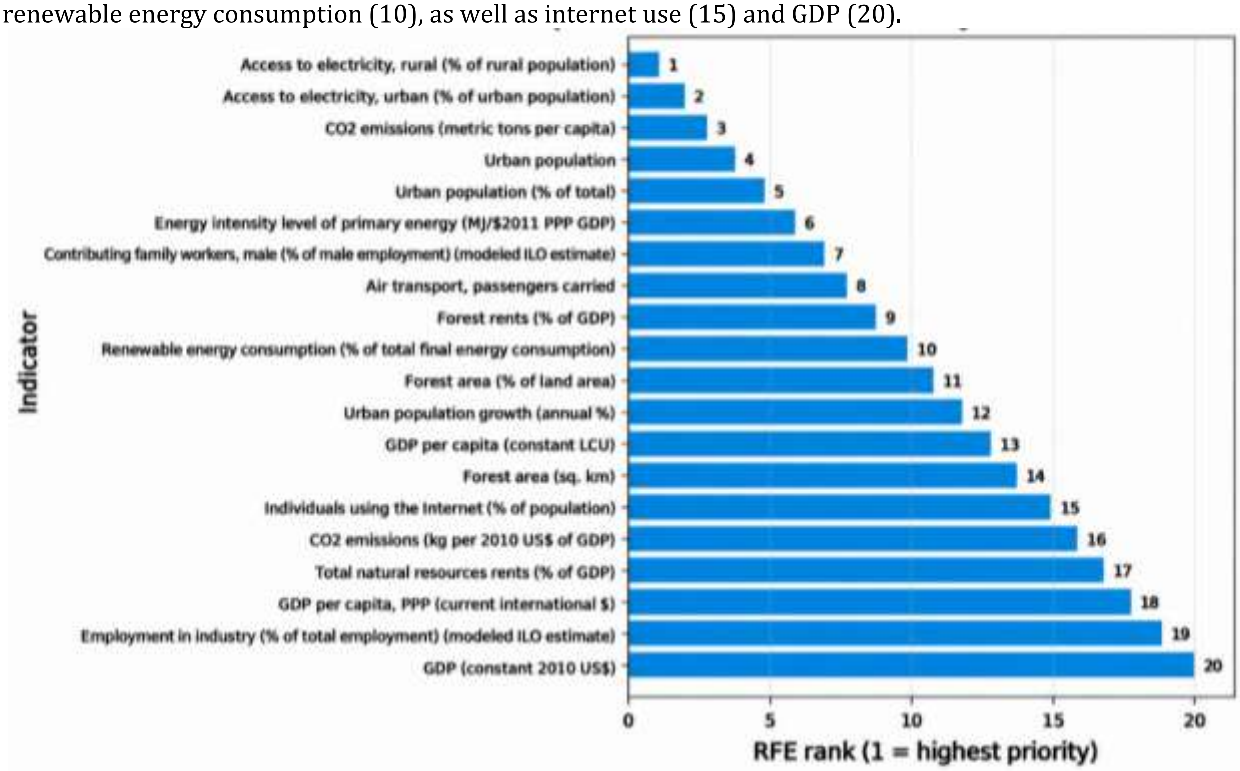

Figure 6: Top 20 Indicators Identified by RFE

Figure 7 shows the LightGBM model's predicted access to electricity (%) versus the actual reported values. The data points closely align with the fit trend line in the observed range, showing strong correspondence between actual and predicted values with low deviation.

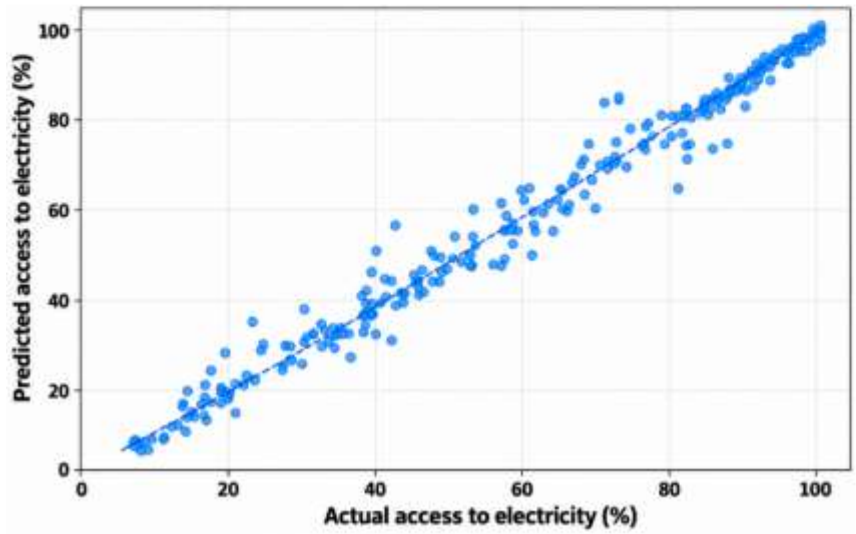

Figure 7: Actual versus Predicted Electricity Access

LightGBM model metrics and evaluation settings are presented in Table 3. The model scored an R^2 of 0.9925, an RMSE of 2.3167, an MAE of 1.0188 and an MAPE of 2.43% on the test data.

Table 3: Performance Evaluation

Evaluation Parameter	Test Result
R^2	0.9925
RMSE	2.3167
MAE	1.0188
MAPE	2.43%

Figure 8 shows the SHAP feature relative contribution of the LightGBM model (mean absolute SHAP values). The highest overall contribution is attributed to urban electricity access (18.783), followed by rural electricity access (5.649) and urban population share (3.294). The remaining indicators show much lower contributions, with energy intensity showing the lowest contribution (0.069).

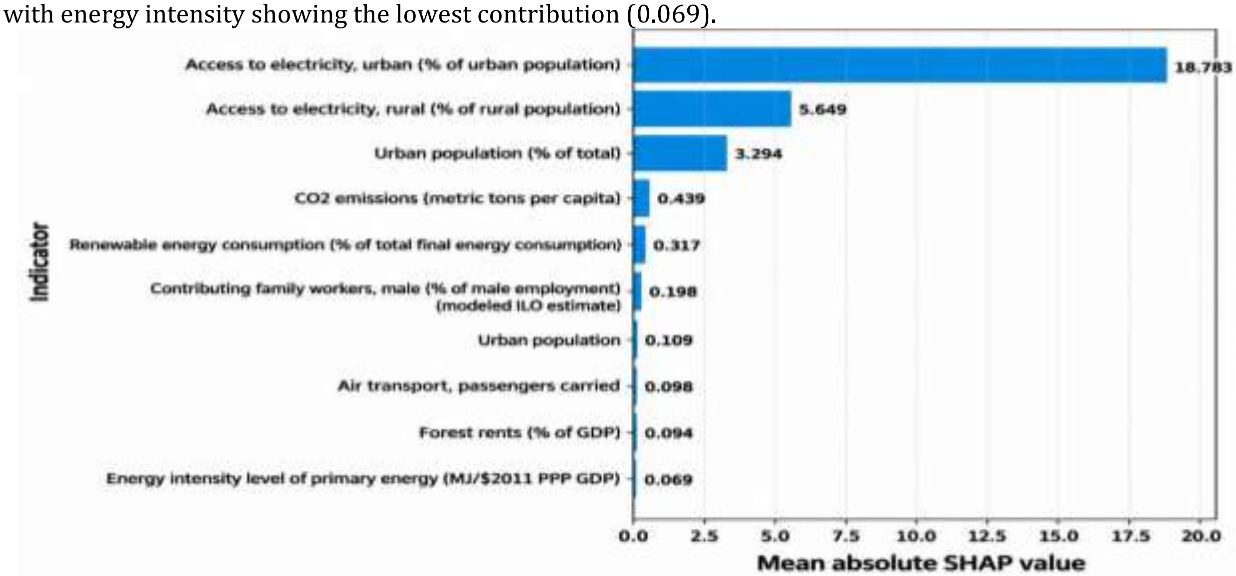

Figure 8: Global SHAP Feature Importance

Figure 9 shows the SHAP contribution plot for LightGBM. The plot shows the contribution of each feature to the model in terms of its importance and direction. The features associated with rural and urban electricity access exhibited the widest SHAP value distributions. Urban population share shows a considerable SHAP value distribution focused around zero. The rest of the features exhibited relatively low SHAP values, with their contributions focused on or around zero.

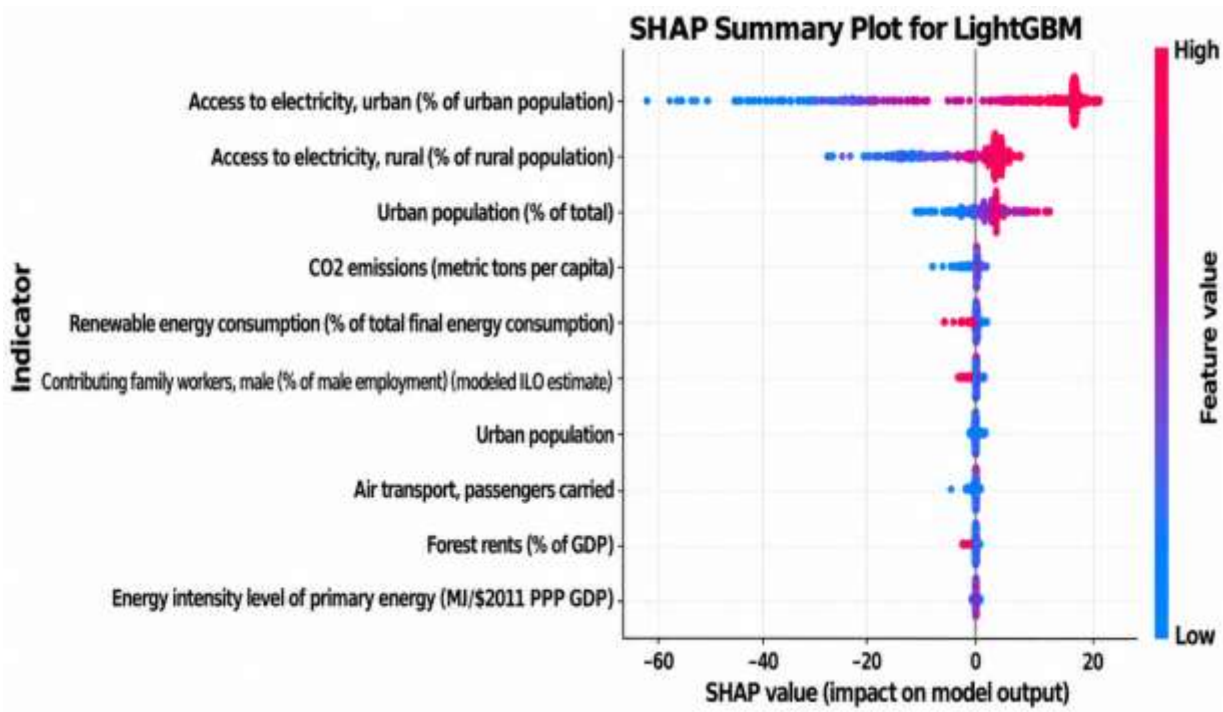

Figure 9: SHAP Summary of Feature Contributions

Figure 10 shows the LOPCOW-based objective weights of the ten assessed criteria in the 2013 baseline year. For this baseline year, the most important criterion is energy intensity (20.79%), followed by CO2 emissions (16.65%), forest rents (15.31%), and urban electricity access (14.29%). The least important criteria, with the lowest weights, are urban population (0.44%) and air transport (0.39%).

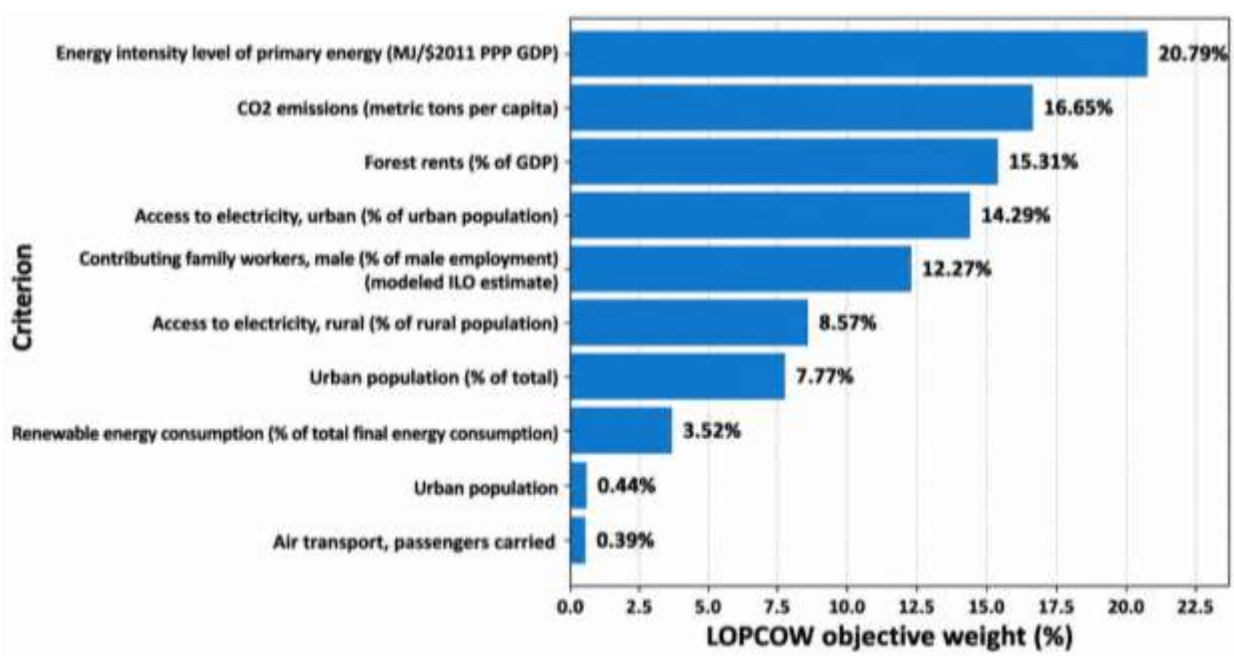

Figure 10: LOPCOW Objective Weights of Selected Criteria

Figure 11 ranks 2013's top 10 entities based on ERUNS's sustainable infrastructure score. Macao SAR, China, ranks first (0.949), followed by Puerto Rico (0.938) and Uruguay (0.932). The other groups have scores between 0.912 and 0.919. This reflects the comparable performance of the top groups.

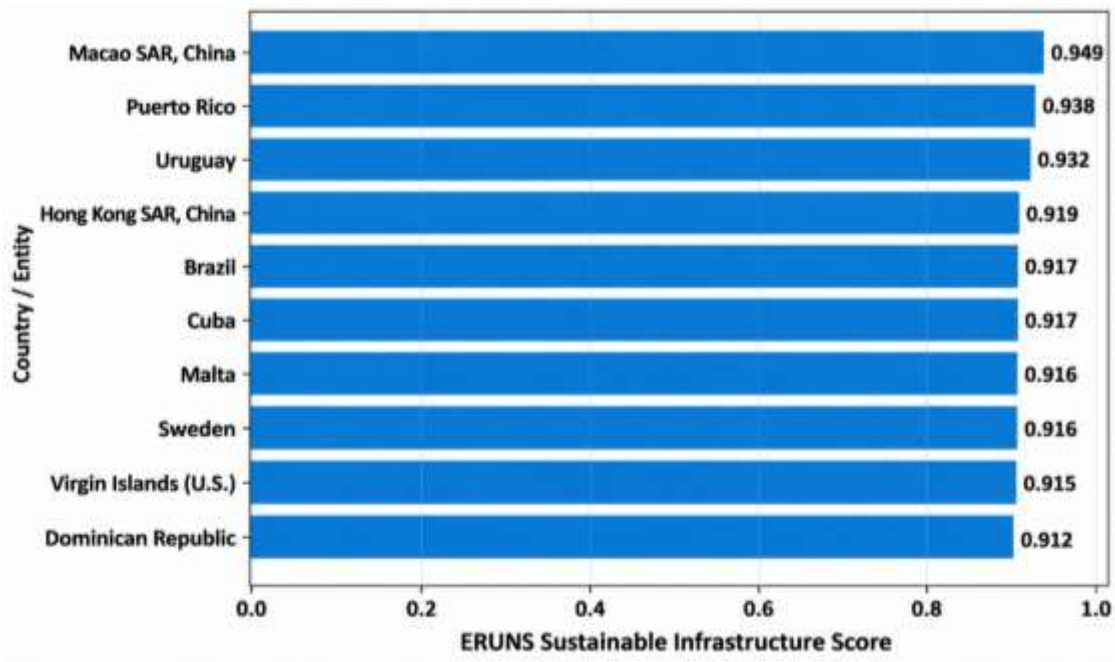

Figure 11: Top 10 Entities by ERUNS Score

Figure 12 shows the validation of ERUNS rankings vs. the TOPSIS benchmark for the year 2013. A strong agreement between the evaluations is shown by the Spearman rank correlation (ρ) of 0.952. With a top 10 overlap of 4/10, note that the four entities commonly identified among the 10 top-ranked entities are: Air Traffic Assistance, Engineering Education, Brain Drain, and Vicinity Tax.

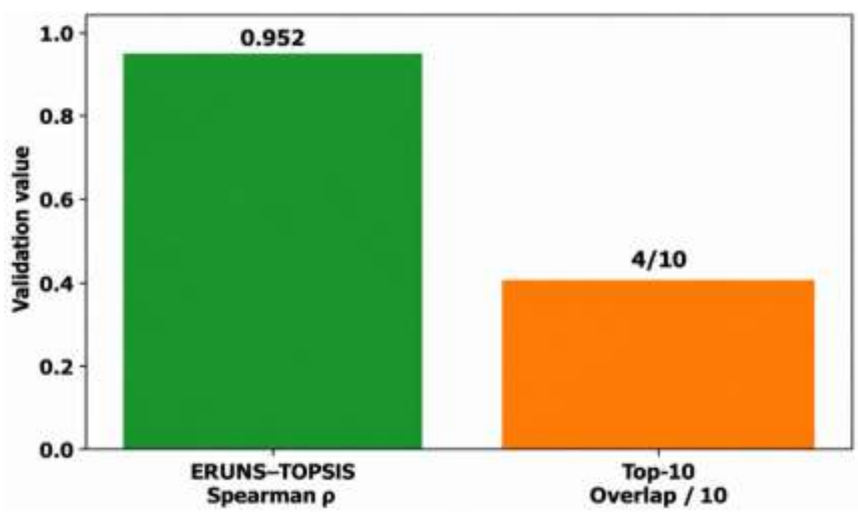

Figure 12: ERUNS Ranking Validation Against TOPSIS

Figure 13 shows the effect of ±10% shifts in each individual criterion weight on the ranking for 2013. Spearman rank correlations show that the stability of the ranking is very high (up to approximately 1.000 for all the criteria) with changes in criterion weights.

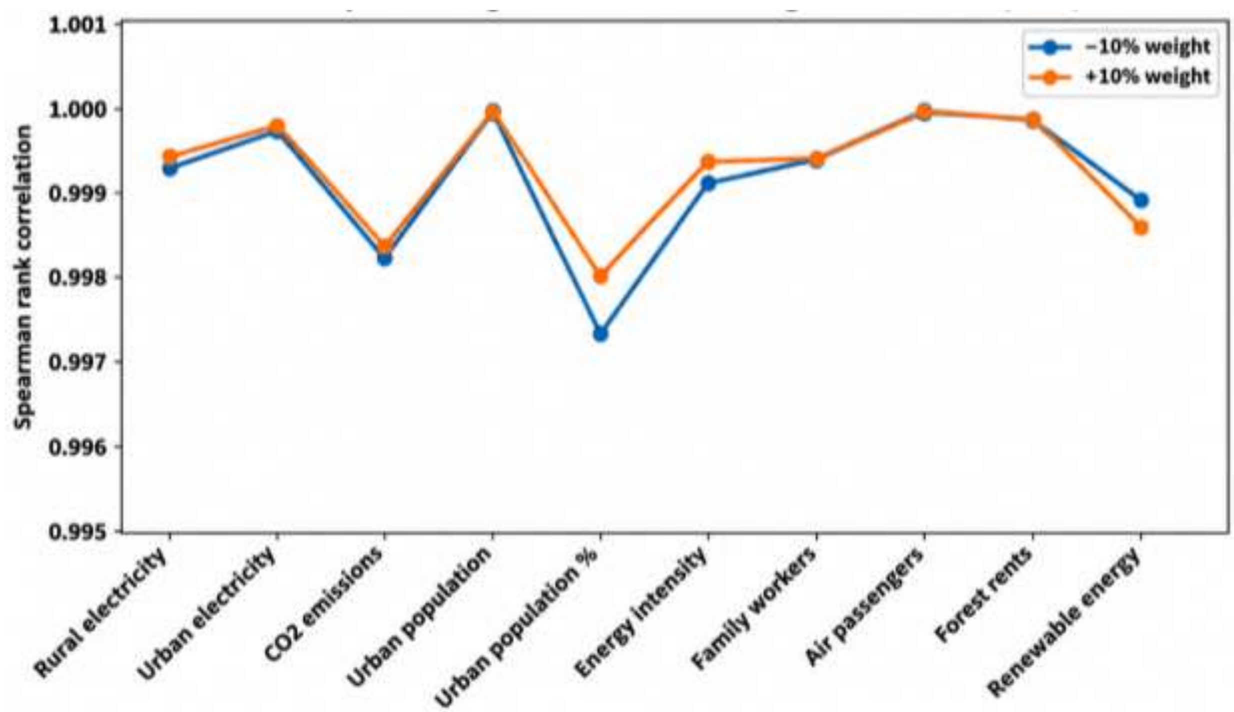

Figure 13: Ranking Sensitivity to Weight Perturbation

4.1 Comparative Analysis

Table 4 shows recent predictive performances of AI-based studies. Studies are shown in reverse chronological order, with the present study shown last. Shohan et al. (2026) reported R^2 = 0.734 with the use of SVM, RF, XGB, and SHAP, while Ghafari and Samaei (2025) published R^2 = 0.910 with a proposed AI-integrated Digital Twin framework. Shen and Pan (2023) reported R^2 = 0.934 using BO-LGBM and SHAP for assessing sustainable buildings. The present study (2026) reported the highest R^2 = 0.9925, with the aid of RFE and LightGBM, and SHAP-based explainability and LOPCOW–ERUNS-based decision analysis. This indicates that the predictive performance of the proposed framework is strong when integrated with explainability, objective weighting, and ranking of sustainable infrastructures.

Table 4: Comparative Analysis

Author	Year	Technique Used	R^2
Shohan et al. [25]	2026	SVM, RF, XGBoost (XGB), SHAP	0.734
Ghafari and Samaei [28]	2025	BIM/GIS, IoT, Remote Sensing, FEM, RF, LSTM	0.910
Shen and Pan [34]	2023	BIM, DesignBuilder, Bayesian Optimization–LightGBM (BO-LGBM), SHAP, AGE-MOEA	0.934
Present Study	**2026**	**RFE, LightGBM, SHAP, LOPCOW, ERUNS**	**0.9925**

5. Conclusion and Future Scope

Assessing the economic, environmental, energy, and social segments of a region for new infrastructure developments requires robust evidence-based approaches for effective planning and resource allocation. Sustainability assessment based on ML and MCDM has gained momentum in recent years; however, there is a lack of integration of predictive modeling, explainability, objective weighting, and ranking of different infrastructure types. The present study integrated the aforementioned gaps through the development of an XAI-based decision support system using global development indicators. The framework incorporated data screening, RFE, LightGBM prediction, SHAP for interpretation, LOPCOW objective weighting, and ERUNS ranking. In RFE, the validation RMSE reached 2.34 with the selected subset of 10 indicators. For the 2014–2018 test period, LightGBM prediction yielded R^2 = 0.9925, RMSE = 2.3167, MAE = 1.0188, and MAPE = 2.43%. SHAP indicated urban (18.783) and rural (5.649) electricity access as the most important predictors. LOPCOW assigned the highest weights to energy intensity (20.79%) and CO_2 emissions (16.65%), and Macao SAR, China, was ranked first with a score of 0.949. The validation Spearman correlation was 0.952, and the TOPSIS ranking was sensitive to a ±10% change in weight. This framework could be extended to incorporate more recent data and alternative objectives, along with additional validation, uncertainty and scenario modeling.

References

1. El-Bouayady, Rachida, Hassan Radoine, Nour-Eddin El Faouzi, Soukaina Tayi, and Hakki Can Ozkan. "Assessing and modeling the impact of urbanization on infrastructure development in Africa: A data-driven approach." Cities 155 (2024): 105486. https://doi.org/10.1016/j.cities.2024.105486
2. Mačiulytė-Šniukienė, Alma, Mindaugas Butkus, and Vida Davidavičienė. "Development of the model to examine the impact of infrastructure on economic growth and convergence." Journal of Business Economics and Management (JBEM) 23, no. 3 (2022): 731-753. https://hdl.handle.net/10419/317578
3. Mishra, Priyanka, and Ghanshyam Singh. "Energy management systems in sustainable smart cities based on the internet of energy: A technical review." Energies 16, no. 19 (2023): 6903. https://doi.org/10.3390/en16196903
4. Mkhitaryan, Khoren, Mariana Kocharyan, Hasmik Harutyunyan, Anna Sanamyan, and Seda Karakhanyan. "Sustainability indicators and urban decision-making: A multi-layer framework for evidence-based urban governance." Urban Science 10, no. 2 (2026): 70. https://doi.org/10.3390/urbansci10020070
5. Gunda, Thushara, Amanda Wachtel, Shruti Khadka Mishra, and Emily Moog. "Quantitative approaches for including equity in risk and resilience infrastructure planning analyses." Risk Analysis 45, no. 11 (2025): 3427-3438. https://doi.org/10.1111/risa.14230
6. Chang, Carlos M., Gianine Tejada Salinas, Teresa Salinas Gamero, Stella Schroeder, Mario A. Vélez Canchanya, and Syeda Lamiya Mahnaz. "An infrastructure management humanistic approach for smart cities development, evolution, and sustainability." Infrastructures 8, no. 9 (2023): 127. https://doi.org/10.3390/infrastructures8090127
7. Alam, Risha, and Kazi Mohammad Khalid Ahsan. "A Meta-Analysis of AI-Driven Geospatial Analytics for Predictive Maintenance of Critical Infrastructure in Developing Economies." International Journal of Scientific Interdisciplinary Research 4, no. 4 (2023): 375-412. https://doi.org/10.63125/rayrex49
8. Paul, Rajesh. "Development of an AI-Driven Predictive Decision Support Model for Multi-Domain Business Analytics and Infrastructure Optimization." American Journal of Interdisciplinary Studies 6, no. 3 (2025): 173-214. https://doi.org/10.63125/ra93qz89
9. https://www.ibef.org/blogs/sustainable-infrastructure-india-s-path-to-a-resilient-future
10. Khoun, Theara, Ate Poortinga, Nyein Soe Thwal, Iván González de Alba, Andrea McMahon, and Carlos Mendez. "Mapping the dimensions of poverty through big data, socioeconomic surveys and machine learning in Cambodia." Social Indicators Research 180, no. 3 (2025): 1593-1618. https://doi.org/10.1007/s11205-025-03718-3
11. Sumon, Md Fakhrul Islam, Md Osiujjaman, Md Azam Khan, Arifur Rahman, Md Kafil Uddin, Laxmi Pant, and Pravakar Debnath. "Environmental and socio-economic impact assessment of renewable energy using machine learning models." Journal of Economics, Finance and Accounting Studies 6, no. 5 (2024): 112-122. https://doi.org/10.32996/jefas.2024.6.5.13
12. Wang, Longfei, Tengfei Long, Wei Jiang, Elhadi Adam, Chunhui Wen, Weili Jiao, and Guojin He. "Economic well-being assessment: a review of traditional and remote sensing approaches." International Journal of Digital Earth 18, no. 1 (2025): 2504137. https://doi.org/10.1080/17538947.2025.2504137
13. Fan, Chao, Jeffrey Xu, Barane Yoga Natarajan, and Ali Mostafavi. "Interpretable machine learning learns complex interactions of urban features to understand socio-economic inequality." Computer-Aided Civil and Infrastructure Engineering 38, no. 14 (2023): 2013-2029. https://doi.org/10.1111/mice.12972
14. Baresi, Umberto, and Alessio Russo. "Can Machine Learning Support Planning for Equitable Green Infrastructure? A Perspective on Opportunities, Risks, and Ethical Pathways." Urban Science 10, no. 7 (2026): 377. https://doi.org/10.3390/urbansci10070377
15. Bellantuono, Loredana, Flaviana Palmisano, Nicola Amoroso, Alfonso Monaco, Vitorocco Peragine, and Roberto Bellotti. "Detecting the socio-economic drivers of confidence in government with eXplainable Artificial Intelligence," Scientific Reports 13, no. 1 (2023): 839. https://doi.org/10.1038/s41598-023-28020-5
16. Choubin, Bahram, Abolfazl Jaafari, Jalal Henareh, Omid Karimi, and Farzaneh Sajedi Hosseini. "Explainable artificial intelligence (XAI) for interpreting predictive models and key variables in flood susceptibility." Results in Engineering 27 (2025): 105976. https://doi.org/10.1016/j.rineng.2025.105976

17. Abdollahi, Arnick, and Biswajeet Pradhan. "Explainable artificial intelligence (XAI) for interpreting the contributing factors feed into the wildfire susceptibility prediction model." Science of the Total Environment 879 (2023): 163004. https://doi.org/10.1016/j.scitotenv.2023.163004
18. Kalasampath, Khushi, K. N. Spoorthi, Sreeparvathy Sajeev, Sahil Sarma Kuppa, Kavya Ajay, and Angulakshmi Maruthamuthu. "A literature review on applications of explainable artificial intelligence (XAI)." IEEE Access 13 (2025): 41111-41140. https://ieeexplore.ieee.org/abstract/document/10908240/
19. D'Adamo, Idiano, Marialucia Della Sciucca, Massimo Gastaldi, and Barbara Lupi. "Indicator assessment of sustainable development goals: a global perspective." Sustainability 17, no. 18 (2025): 8259. https://doi.org/10.3390/su17188259
20. Eşiyok, Sevgi, Erhan Ariş, and Figen Antmen. "Ranking and evaluation of G7 countries and Turkey by gggi indicators using ENTROPY, CRITIC and EDAS methods." Çukurova Üniversitesi Mühendislik Fakültesi Dergisi 38, no. 3 (2023): 647-660. https://dergipark.org.tr/tr/pub/cukurovaumfd/article/1377228?issue_id=80473
21. Lin, Runze, Xinbei Lv, Kunyang Li, Xinning Lv, Zhiwei Zhou, Yuxiao Kou, and Jizhong Shao. "Assessing the renewal urgency of aging residential communities using explainable machine learning: A case study of Wuhan's central urban area." Available at SSRN 7082931. http://dx.doi.org/10.2139/ssrn.7082931
22. UZUNALİ, Şeyma YİĞİT. "Spatial Modeling of Anthropogenic Forest Fire Risk from a Landscape Planning Perspective: Comparison of Expert-Based MCDM and Machine Learning Approaches in Kahramanmaraş, Türkiye." MAS Journal of Applied Sciences (2026). file:///C:/Users/PC/Downloads/14+numara+%C5%9Feyma_pagenumber%20(1).pdf
23. Haque, BM Taslimul, and Md Arifur Rahman. "Data Preparation and Feature Engineering in High-Dimensional Predictive Modeling." Journal of Sustainable Development and Policy 3, no. 04 (2024): 205-244. https://doi.org/10.63125/5txtr530
24. Anderson, Michael, Jennifer Collins, David Thompson, Emily Carter, and Charles Paul. "AI-Based Decision Support Systems for Infrastructure Deployment and Cost Efficiency." https://www.researchgate.net/profile/Charles-Paul-8/publication/404152382
25. Shohan, Ahmed Ali A., Mohammed Alshayeb, and Saleh Alsulamy. "Predicting sustainability performance in construction projects using machine learning: a comparative study." Scientific Reports 16, no. 1 (2026): 2828. https://doi.org/10.1038/s41598-025-32658-8
26. Wang, Zi, and Fang Ren. "Developing a decision support system for sustainable urban planning using machine learning-based scenario modeling." Scientific Reports 15, no. 1 (2025): 13210. https://doi.org/10.1038/s41598-025-90057-5
27. Fedorka, Pavlo, Roman Buchuk, Mykhailo Klymenko, Fedir Saibert, and Andrii Petrushyn. "The use of adaptive artificial intelligence (AI) learning models in decision support systems for smart regions." J. Res. Innov. Technol 4, no. 1 (2025): 99-115. https://www.ceeol.com/search/article-detail?id=1341463
28. Ghafari, Rasoul, and Seyed Reza Samaei. "An AI-driven and digital twin-based framework for climate-resilient and sustainable infrastructure management in Palm Jumeirah: A case study integrating remote sensing." In Proceedings of the Second Conference on Civil Engineering, Architecture, Urban Planning, and Environment, vol. 2, no. 2, 2025. https://www.researchgate.net/profile/Rasoul-Ghafari/publication/390535965_An_AI-Driven_and_Digital_Twin-Based_Framework_for_Climate
29. Lutfiani, Ninda, Nuke Puji Lestari Santoso, Ridhuan Ahsanitaqwim, Untung Rahardja, and Achani Rahmania Az Zahra. "AI-based strategies to improve resource efficiency in urban infrastructure." International Transactions on Artificial Intelligence 2, no. 2 (2024): 121-127. https://doi.org/10.33050/italic.v2i2.545
30. Shulajkovska, Miljana, Maj Smerkol, Gjorgji Noveski, and Matjaž Gams. "Enhancing urban sustainability: Developing an open-source AI framework for smart cities." Smart Cities 7, no. 5 (2024): 2670-2701. https://doi.org/10.3390/smartcities7050104
31. Hadiyana, Thomas, and Seo Ji-hoon. "AI-driven urban planning: Enhancing efficiency and sustainability in smart cities." ITEJ (Information Technology Engineering Journals) 9, no. 1 (2024): 23-35. file:///C:/Users/PC/Downloads/aguspamuji,+3+Enhancing+Efficiency+and+Sustainability+in+Smart+Cities%20(4).pdf
32. Tomar, Deepak. "AI-Based Decision Support Systems for Sustainable Development Policy and Planning." Development 3 (2024): 4. https://doi.org/10.55041/IJSREM29769
33. https://www.kaggle.com/datasets/theworldbank/sustainable-development-goals
34. Shen, Yuxuan, and Yue Pan. "BIM-supported automatic energy performance analysis for green building design using explainable machine learning and multi-objective optimization." Applied Energy 333 (2023): 120575. https://doi.org/10.1016/j.apenergy.2022.120575