International Journal of Artificial Intelligence and Machine Learning
Publisher's Home Page: https://www.svedbergopen.com/
Research Paper Open Access

Sustainable Infrastructure Systems For Green Buildings: Lifecycle Optimization And Performance Modeling Using Machine Learning

Bipin Sule[1], Jyoti Mahur[2], Zhou Yiting[3], Tanya Singh[4], Kanchana K[5], Govind Singh Panwar[6], L. Sathiya[7], Monisha J[8], Rahul Rajendra Papalkar[9]

[1] Department of Engineering, Science and Humanities, Vishwakarma Institute of Technology, Pune, Maharashtra 411037, India. Email: bipin.sule@vit.edu
[2] Assistant Professor, Department of School of Engineering & Technology, Noida International University, India. Email: jyoti.mahur@niu.edu.in
[3] Research Scholar, School of Business and Management, Lincoln University College, Malaysia. Email: yiting.phdscholar@lincoln.edu.my
[4] Professor, Department of School of Engineering & Technology, Noida International University, India. Email: dean.academics@niu.edu.in
[5] Assistant Professor, Department of Commerce, Meenakshi College of Arts and Science, Meenakshi Academy of Higher Education and Research, India. Email: kanchana@maher.ac.in
[6] Assistant Professor, Department of Computer Science & Engineering, Dev Bhoomi Uttarakhand University, Dehradun, Uttarakhand, India. Email: pc.cse@dbuu.ac.in ORCID: 0009-0007-5602-1512
[7] Assistant Professor, Department of CSE, Panimalar Engineering College, Tamil Nadu, India. Email: sathiyamani@gmail.com ORCID: 0009-0004-5166-9393
[8] Assistant Professor, Department of Management Studies, Meenakshi College of Arts and Science, Meenakshi Academy of Higher Education and Research, India. Email: jmonishamba@maher.ac.in
[9] Department of Computer Engineering, Vishwakarma University, Pune, Maharashtra 411048, India. Email: rahul.papalkar@vupune.ac.in

Abstract

Resource-efficient designs, lifecycle assessment, and intelligent control mechanisms characterize sustainable infrastructures that minimize the environmental impacts while maintaining functionality. Nevertheless, current methodologies are unable to optimize the lifecycle process, predict future events, and incorporate sustainability objectives. This research focuses on developing an innovative machine learning (ML) system for lifecycle optimization and performance modeling of green buildings. The research aims to promote sustainability, energy efficiency, and occupant comfort in green buildings. The Green Build Lifecycle Performance data consists of 9,400 entries having 40 attributes is adopted for this research. Energy consumption, environmental characteristics, occupancy behaviors, and other sustainability variables are recorded in the dataset. Data pre-processing is performed through Min-Max normalization, and the most important features are extracted using Principal Component Analysis (PCA). The dataset is employed to train predictive models and improve decision-making throughout the building lifecycle. An Efficient Ant Colony-tuned Dynamic Random Forest Classifier (EAC-DRFC) is proposed for predictive models. EAC helps select features and tune hyperparameters for the EAC-DRFC model, whereas DRFC helps learn adaptively based on the dynamic nature of the data. The proposed model using Python is impressive in terms of its accuracy of 0.97, precision of 0.96, recall of 0.98, and F1-score of 0.97. The research offers intelligent, adaptive, and lifecycle-oriented management of sustainable green building infrastructure systems.

Keywords: Sustainable Infrastructure, Green Buildings, Lifecycle Optimization, Performance Modeling, Energy Efficiency, Environmental Sustainability.

1. Introduction

Environmental issues, social problems, and economic constraints that affect the construction industry are described using an idea referred to as green buildings. Sustainable infrastructural design has a significant part to play in emission reduction, efficiency improvement, and urban sustainability through integration of resource efficiency, resource management, and energy efficiency into the building [1]. Green building management systems use cyber-physical technology to enhance sustainability and efficiency. Nevertheless, as there are various cybersecurity threats that potentially disrupt building management services, lead to a breach of sensitive data, and make the achievement of sustainability objectives impossible, it is necessary to ensure both the security and resilience of such infrastructure [2]. he quantity of energy, water, and other resources required for constructing a building can be highly detrimental to the natural environment. The primary objective is to maximize the efficiency of resource usage; minimize waste, risks, and emissions to increase the sustainability of the infrastructure and construction process; and promote resiliency [3]. The construction of buildings consumes many different resources. As a result, one could say that it makes a considerable contribution to environmental degradation. The primary objective of sustainable infrastructure and construction processes is resource utilization and efficiency [4]. The importance of the integration process for green infrastructure can be illustrated by its implications on the environment, particularly urban heat islands due to increased urbanization. Nonetheless, researching the various services offered by green infrastructure and the measurement process is challenging due to a lack of sufficient knowledge on the subject [5, 6]. Sustainable infrastructure needs proper management with respect to supply chain, energy, and the environment. As a means of achieving sustainability in many projects, which explore green financing and green technology innovations. Survey-based and impact analysis-based methods are the commonly used methodologies for conventional approaches; nevertheless, these have some limitations, including bias when gathering data, insufficient foresight, and several others [7]. Traditional approaches for assessing the environmental impact of green buildings are surveys, case studies, and life cycle assessment (LCA). Nonetheless, there are some challenges in using these methodologies, such as data validity issues, high costs, and insufficient knowledge on sustainable practices, among others [8].

Research Aim: The research utilizes ML algorithms to evaluate sustainability across environmental, economic, and social dimensions in green buildings, addressing the limitations of traditional analysis methods. The proposed EAC-DRFC method enhances this process by performing optimal feature selection and hyperparameter tuning through EAC, while DRFC adaptively learns from the dynamic characteristics of the data. By estimating the impacts of energy consumption, carbon emissions, and overall sustainability performance, the proposed approach enables more accurate analysis and supports informed decision-making throughout the lifecycle of green buildings.

Research Organization: The structure framework of the research is listed as follows Application of the ML Algorithm for Evaluating Sustainability of Green Buildings are presented in Section 1. The gaps in current knowledge and their analysis of green buildings through existing approaches are discussed in Section 2. The methodology of the suggested algorithm for the assessment of the influence of green buildings on the sustainability of the environment and economy is introduced in Section 3. The testing results and discussion are covered in Section 4. The rest of the research is presented in Section 5.

2. Related Works

Analysis of the current literature review on sustainable green buildings, which focuses on objectives, methods, outcomes, and limitations associated with intelligent infrastructures, technological advances in green technologies, sustainable urban development programs, and applications of digital innovations, is provided in Table 1 below.

Table 1: Comparative analysis of sustainable green building research

Ref	Objective	Method	Result	Limitations
[9]	Evaluate the impact of green buildings on health and indoor environmental quality (IEQ).	Assessment of sustainable construction and scientific literature	Green buildings improve IEQ and health conditions.	Health implications neglected
[10]	Assess the role played by green structures in	Consider examples and discuss the benefits and	Green buildings enhance resiliency, energy usage, and	Legislation, public perception, and monetary constraints all hinder implementation.

	ensuring resiliency in urban areas.	challenges of green building.	health in urban centers.	
[11]	Analyze developments in Green Building Technologies (GBT) for mitigating climate change and promoting sustainable development.	Analyze current research using a methodical manner to identify GBT issues.	Development of ecologically acceptable and commercially viable solutions.	Insufficient research on strategic advancement planning.
[12]	Analyze intelligent green buildings' digital advancements.	Intelligent Green Buildings (IGB) concepts and the integration of Digital Twins (DT).	IGB improved sustainable urban development, and DT enhanced sustainability strategies.	Insufficient incorporation of DT into intelligent building designs.
[13]	Investigate the use of green building materials.	Observation and sampling methods.	Adoption of green materials was highly insufficient.	High cost and low market demand.
[14]	Combine a culture of communal well-being with sustainability.	Participatory design and sustainability examination.	Green design improved environmental and social well-being.	Difficulty integrating cultural and social concerns with sustainability.
[15]	Evaluate how effectively green technology reduces urban heat island effects.	Simulation of green walls and roofs design Builder.	Green technologies lowered cooling load requirements and conserved energy.	Research focused only on hot and humid regions.
[16]	Develop predictive models for green building design optimization.	Exploratory Data Analysis (EDA) and Machine Learning ML models.	Traditional ML and Deep learning (DL) perform better in accuracy.	Increased computational complexity.

2.1. Research gap: Existing research methods, life-cycle assessment, and conventional ML algorithms form the core of modern research on sustainable green buildings. All these conventional methods often face issues related to data bias, computational complexity, adaptability issues, and failure to consider real-world dynamics. These conventional approaches are inconsistent in making predictions and have been known to encounter difficulties in working with high-dimensional data. The proposed approach of EAC-DRFC bridges all these gaps by promoting improved decision-making and node splitting capabilities, as well as feature selection.

3. Methodology

Figure 1 is an ML paradigm that illustrates the process of developing sustainable infrastructure systems in green buildings, from data collection/preprocessing to feature selection via min-max normalization, principal component analysis (PCA), and optimization by EAC-DRFC. The evaluation process focuses on accuracy, precision, recall, and F1-score.

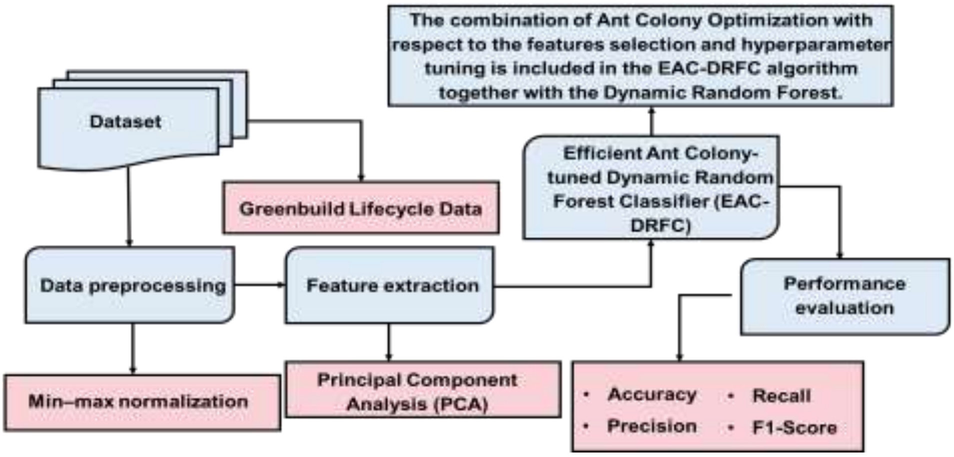
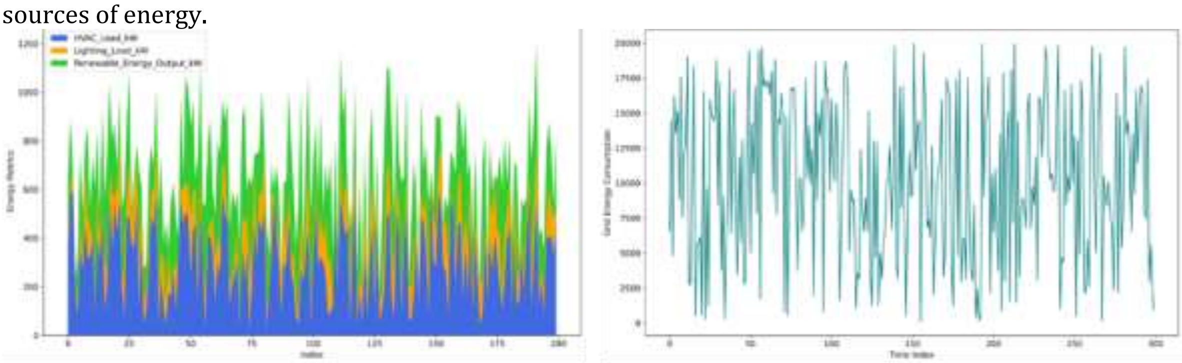


Figure 1: EAC-DRFC Framework for Sustainable Green Buildings

3.1 Data collection

The Green Build Lifecycle Performance database has 9,400 records that comprise 40 attributes that indicate sustainable infrastructural design and green building performance assessment. The attributes include a number of features ranging from energy consumption, environmental impacts, occupancy, comfort level, lifecycle parameters such as HVAC load, production of renewable energy, carbon footprint, indoor air quality, and resource consumption. This data set comprises 80% training data and 20% test data sets, enabling prediction of sustainability. **Kaggle Source:**
https://www.kaggle.com/datasets/colabsss/greenbuild-lifecycle-data

Data feature exploration: Figure 2(a) is utilized to illustrate the loads of HAVC systems, lighting, and renewable energy production in green buildings. The figure provides insight into how energy balance and sustainability of green buildings can be assessed using the presented data. Figure 2(b) illustrates the variations in grid energy consumption over time. This implies that the buildings are dependent on external sources of energy.



(a) (b)

Figure 2: Energy performance indicators and grid consumption (a) Green Building Energy Production and Consumption Trends, (b) Variation of Grid Energy Consumption with Time

3.2 Data preprocessing using min–max normalization

The Min-Max Normalization procedure is employed to scale different kinds of green building data to ensure that none of the green building sustainability factors dominate other factors in the process of training. As stated in accordance with establishing an intelligent sustainable infrastructure, this method allows for normalizing such variables as energy consumption, temperature, humidity, and air quality, ensuring that the learning process is stable and accurate. The widely applied technique in this case involves mapping feature values into a certain range, usually [0-1].

$$U^p = \overline{max_B - min_B}^{u-min} B\left[new_{max_B} - new_{min_B}\right] + new_{min_B} \quad (1)$$

Equation (1) above is a version of the min-max normalization equation adapted to fit sustainable green building data. In this context, U^p stands for the normalization of any sustainability parameter u including the energy consumption, temperature, or air pollution level. The variables $max_B - min_B$or the maximum and minimum values of the variables in dataset B. This shows the span of the variables in the set of building parameters. The equation $\overline{max_B - min_B}^{u-min}$ shifts the data along the axis until its minimum value equals zero. Subsequently, dividing by $max_B - min_B^{-}$ scales the variable to the original span. The final step involves multiplying the data by the equation $new_{max_B} - new_{min_B}$. The purpose of this is to scale the data into the required span, usually [0,1].

3.3 Feature extraction using Principal Component Analysis (PCA)

The heterogeneity level of the green building data set is the highest when the PCA approach is used, which is one method of linear transformation used in mapping higher dimensions of sustainable property values to lower dimensions. PCA is an optimal technique used in speeding up processing and adaptive life cycle performance simulation, which is easy to process and ensures reduced computing time while discarding environmental variables that are not required for creating intelligent and energy-efficient infrastructure systems. The main reason why PCA components are orthogonal and independent is that they come from eigenvectors of the covariance matrix, which contains the most important sustainability trends of the data set.

$$T = \frac{1}{l}\sum_{j=1}^{l}(w_j - \mu)(w_j - \mu)^s \qquad (2)$$

In Equation (2), T is the covariance matrix used to identify sustainability features in green building systems. In this, l relates to the sample size, which converts the total into an average, $\sum$ (sigma) is the used to represent summation in mathematics. j refers to each data point's indexing, which begins at 1 and ends at l. The term $(w_j - \mu)$ represents the difference between the average values of the sustainability factors (temperature, energy, and air quality) value μ. The expression $(w_j - \mu)(w_j - \mu)^s$ determines the outer product, which shows how the variables are related.

$$z_l = b_{l1w1} + b_{l2w2} + \cdots + b_{llwl} \qquad (3)$$

In Equation (3), z_l appears to be a linear function of the vectors of sustainability characteristics$w_1, w_2, \ldots, w_l$.. Features related to buildings like energy usage, temperature, humidity, and air quality are all captured by these vectors, as well as the product of each of these feature vectors by a weight $b_{l1}, b_{l2}, \ldots, b_{ll}$. This is a way of showing how crucial that particular characteristic is in modeling green building behavior. As a result, one can obtain z_l.

$$\gamma_l = \frac{\lambda_1+\lambda_2+\cdots+\lambda_n}{\lambda_1+\lambda_2+\cdots+\lambda_n+\cdots+\lambda_l} \geq 80\% \qquad (4)$$

In Equation (4), $\gamma_l$represents the proportion of variance of the respective sustainability parameter in the data related to green buildings $\lambda_1, \lambda_2, \ldots, \lambda_n$ to the sum of all eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_l$. It is widely applied in PCA for calculating the percentage of variance captured by the chosen principal components. Every eigenvalue λ_l stands for the proportion of variance for that specific sustainability element in relation to information about green buildings, such as energy usage, environmental factors, and occupant comfort. When $\gamma_l \geq 80\%$, it implies that the chosen components preserve at least 80% of the total variance.

3.4 Sustainable Infrastructure Systems for Green Buildings using EAC-DRFC

Proposed approach for sustainable infrastructure systems in green buildings through an ML-based system with EAC and DRFC. The EAC algorithm is employed to help with reducing dimensions, increasing computational efficiency, and enhancing the relevance of the models to variables associated with sustainability. The DRFC is incorporated to ensure the adaptability of the tree structures in response to data changes. These techniques are instrumental in improving energy efficiency analysis and prediction, ensuring effective decision-making for sustainable green buildings.

DRFC for Prediction Accuracy: The research applies a DRFC approach to enhance sustainability prediction in green building infrastructure systems. While traditional Random Forests offer robustness, in conventional RF algorithms, the criterion used to split nodes remains static, thereby restricting flexibility in adapting to changing patterns of data, limiting adaptability, causing poor performance when using high-dimensional attributes, and ultimately leading to lower prediction accuracy in sustainability assessment scenarios. In the suggested DRFC method, adaptive selection of nodes and composite decision rules increase classification consistency, efficiently manage multidimensional data, and improve prediction accuracy for sustainability evaluation.

$$Gain(C,b) = Ent(C) - \sum_{u=1}^{U}\frac{|C^u|}{|C|}Ent(C^u) \qquad (5)$$

In Equation (5), $Gain$ evaluates the efficiency of the attribute which increases accuracy in decision making and prediction in the green building dataset, C the complete set (green building dataset: energy, occupancy, environment), b a single feature (temperature, energy consumption, etc.), $Ent(C)$ entropy of the whole set, $\sum_{u=1}^{U}$ implies summation of effects of all partitions formed through b, $\frac{|C^u|}{|C|}$ the weight of each partition, $Ent(C^u)$ entropy after dividing the whole set through feature b.

$$Gini(C) = 1 - \sum_{l=1}^{|z|} ol^2 \qquad (6)$$

The Equation (6) below demonstrates how the $Gini$ index represents the segregation of data that allows the model to make accurate predictions regarding the green building performance, where $|z|$ represents the number of classes (for instance, performance grade) while l denotes the index of the label/class and ol^2 represents the squared probability of the class l.

$$G = \min_{\alpha,\beta \in Q} E\{C, b\} = \alpha Gini(C, b) - \beta Gain(C, b) \quad (7)$$

In Equation (7), G is the last optimized value which helps to get the most optimal feature selection to predict green building performance efficiently, $\min_{\alpha,\beta \in Q}$ optimizes the weights to minimize the objective function for improved prediction in the green building model, $E\{C, b\}$ represents the objective score that selects the best set of features to make accurate predictions about sustainable green building performance, $\alpha Gini(C, b)$ decides the importance placed by the model on reducing the impurity of data during accurate green building performance classification, $\beta Gain(C, b)$ decides the importance placed by the model on informative feature selection for improved sustainability prediction in green building systems.

EAC for Optimal Hyperparameter Tuning: The proposed method uses an advanced EAC to improve feature selection and hyperparameter tuning for sustainable green building systems. Traditional ant colony algorithms suffered from premature convergence and local optima, limiting global search in high-dimensional datasets. The EAC introduces adaptive pheromone updates, dynamic evaporation rates, and a balanced exploration–exploitation strategy, enhancing optimization efficiency and predictive performance.

$$\tau_{ji}(s+1) = \begin{cases} \rho^{\int_s^{s+1} \frac{\tau_{ji}(s) T_{min}}{T_{max}}} \tau_{ji}(s), \tau \geq \gamma_{max} \\ \rho^{\int_s^{s+1} \frac{\tau_{ji}(s) T_{max}}{T_{min}}} \tau_{ji}(s), \tau < \gamma_{min} \end{cases} \quad (8)$$

In Equation (8), τ (pheromone value) is the importance weighting applied to the selection of choosing a feature or decision tree from j to i, s denotes the stage within the EAC procedure. $(s + 1)$ is the subsequent stage following the application of the pheromone values, and ρ is the evaporation coefficient for pheromones, which encourages the algorithm to be flexible enough to accommodate changes in building conditions. $\int$ describes changes occurring during an optimization step, τ_{ji} represents the importance of a feature/path in making predictions at the current stage, T represents the quality of a chosen set of features or model settings, T_{min} represents the minimum acceptable level of system performance, T_{max} represents the optimal (maximum) value of system performance. $\tau \geq \gamma_{max}$ indicates the presence of significant features and implements controlled reinforcement to ensure reliable and consistent predictions of sustainability. $\tau < \gamma_{min}$ highlights the unimportant attributes and assists in exploring new ways to find more successful attribute combinations for forecasting sustainable building performance.

$$\Delta\tau_{ji}(s) = \begin{cases} \frac{R}{T_{max}}, X(j,i) \in T_{max} \\ \frac{R}{T_{min}}, X(j,i) \in T_{min} \\ 0, Others \end{cases} \quad (9)$$

In Equation (9), $\Delta\tau_{ji}$ represents the increment in importance due to transition from j to i along path s, R refers a constant reward or performance-based rewards, and denotes the quality of the solution, $X(j,i)$ denotes a feature from node j to node i for EAC. $X(j,i) \in T_{max}$ link belongs to the high-performance group, $X(j,i) \in T_{min}$ link belongs to the low-performance group.

Algorithm 1: EAC-DRFC for Sustainable Lifecycle Optimization in Green Buildings

$Input$: $Dataset\ D = \{energy, HVAC, air\ quality, occupancy, carbon\ emission\}$
$Output$: $Optimized\ sustainability\ prediction\ model$
1: $Load\ green\ building\ dataset\ D$
2: $Normalize\ data\ using\ Min - Max\ normalization$
$\quad U\hat{}p = (u - min\_B) / (max\_B - min\_B)$
3: $Apply\ PCA\ for\ feature\ extraction$
$\quad T = covariance\ matrix$
$\quad z\_l = transformed\ sustainability\ features$
4: $Split\ dataset\ into\ training\ and\ testing\ sets$
5: $Initialize\ DRFC$
$\quad - Number\ of\ trees\ T, Entropy\ Ent(C), Gini\ index\ Gini(C)$
6: $Train\ initial\ DRFC\ model$
7: $Compute\ feature\ selection\ score$
$\quad Gain(C, b)\ and\ Gini(C)$

8: $Initialize\ EAC$
$-\ Pheromone\ \tau\_ji, Evaporation\ rate\ \rho, Reward\ factor\ R$
9: $Generate\ feature\ subsets\ using\ ants$
10: $Update\ pheromone\ values$
$\tau\_ji(s+1)$
11: $Select\ optimal\ features\ and\ hyperparameters$
12: $Retrain\ DRFC\ using\ selected\ features$
13: $Dynamically\ update\ tree\ structures$
14: $Compute\ fitness\ function$
$G\ =\ \alpha Gini(C,b)\ -\ \beta Gain(C,b)$
15: $Compare\ F\_new\ with\ F\_old$
16: $Update\ best\ solution\ if\ F\_new\ >\ F\_old$
17: $Repeat\ optimization\ until\ maximum\ iterations\ reached$
18: $Train\ final\ EAC - DRFC\ model\ and\ predict\ sustainability\ performance$
$Return$: $Sustainable\ lifecycle\ optimization\ and\ prediction\ results$

Algorithm 1 shows how the EAC-DRFC is proposed to efficiently optimize the lifecycle of green buildings using the EAC and DRFC. The EAC-DRFC starts with the building datasets being preprocessed using Min-Max normalization. It also filters out important sustainability properties with PCA. The hyperparameters are tuned and subsets of features are selected using EAC. The DRFC, tunes the ensembles to the distribution of data throughout the time.

4. Result and Discussion

The proposed EAC-DRFC framework has been evaluated and demonstrated improved prediction accuracy, adaptive learning capability, and enhanced sustainability assessment performance in green building infrastructure systems. The algorithm is developed by using python programming and runs on the windows environment on an Intel® core™ i7 processor.

Accuracy: The proportion of accuracy in determining the sustainable infrastructure conditions is what determines precision. To enable minimizing errors in sustainable classification within green building infrastructure systems and ensure dependability in the environmental evaluation process, precision is crucial. P**recision:** The system's accuracy is determined by how effectively it can recognize the overall circumstances of environmentally balanced and energy-efficient green building infrastructure. The accuracy of sustainability forecasts reflects their correctness. **Recall:** During the various stages of the development of green building infrastructure, recall measures are employed to find all pertinent sustainable infrastructure patterns that are essential for effectively differentiating environmental performance. F**1-Score:** The combines precision and recall, guarantees sustainability evaluations and precise forecasts for green building infrastructure.

Performance evaluation using existing dataset [16]: The Proposed model training and evaluation on the existing sustainable infrastructure dataset [16], which has characteristics such as environmental factors, energy usage, and building performance. From the comparison shown, the proposed model performs better in prediction than the traditional methods, such as random forest (RF), decision tree (DT), XGBoost, and stacking [16]. This is supported by the increased performance indices shown in Table 2 and Figure 3.

Table 2: Comparative Sustainable Infrastructure Prediction Models

Models	Accuracy	Precision	Recall	F1-score
RF [16]	0.84	0.82	0.84	0.80
DT [16]	0.76	0.77	0.76	0.77
XGB [16]	0.84	0.83	0.84	0.81
Stacking [16]	0.84	0.81	0.84	0.80
EAC-DRFC [proposed]	0.95	0.92	0.94	0.96

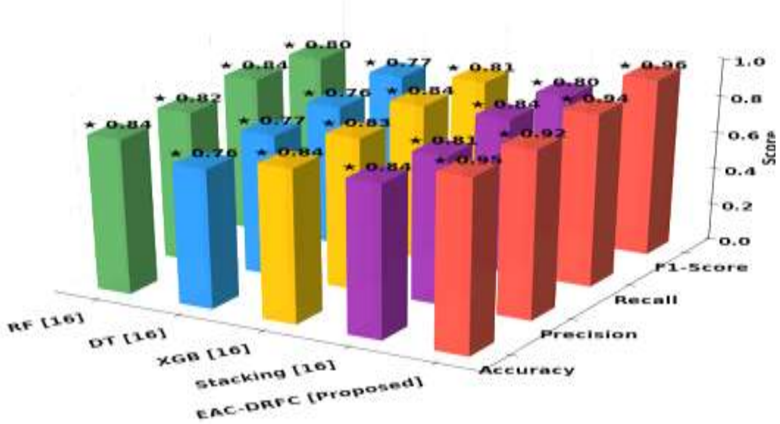


Figure 3: Performance Evaluation of Proposed EAC-DRFC Model

Performance evaluation using the proposed dataset: The existing models, such as RF, DT, XGBoost, and stacking [16], were trained using the proposed datasets of Green Build Lifecycle Data, which contained both structured numeric and temporal features. Table 3 and Figure 4(a-d) show that the proposed EAC-DRFC algorithm is more efficient than the other algorithms with an accuracy of 0.97, a precision of 0.96, a recall of 0.98, and an F1 score of 0.97.

Table 3: Performance Evaluation of Proposed EAC-DRFC Model

Models	Accuracy	Precision	Recall	F1-score
RF	0.92	0.91	0.93	0.91
DT	0.89	0.87	0.90	0.91
XGB	0.92	0.93	0.94	0.92
Stacking	0.93	0.95	0.96	0.97
EAC-DRFC [proposed]	0.97	0.96	0.98	0.97

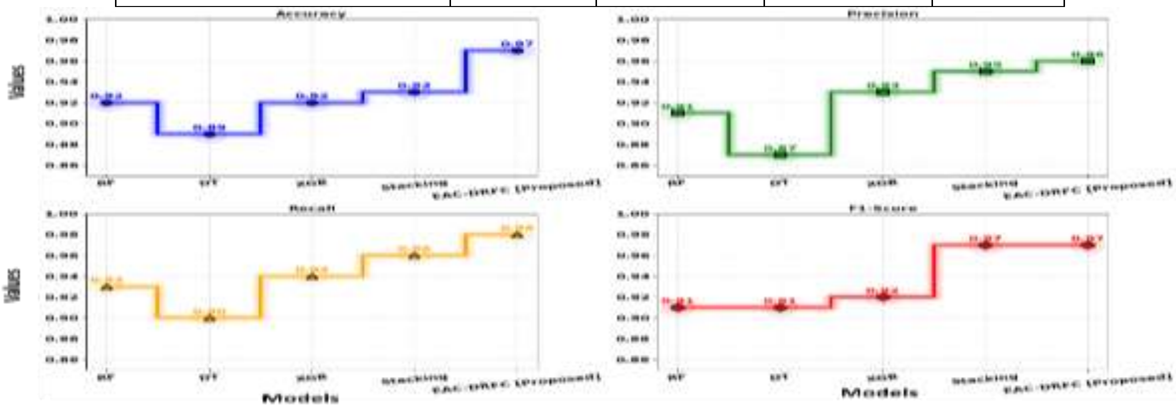


Figure 4: Graphical illustration of (a) accuracy, (b) Precision, (c) Recall, and (d) F1-score.

Existing approaches that include random forest (RF), decision tree (DT), extreme gradient boosting (XGBoost), and ensemble are claimed to be less flexible and inefficient for use in the high dimensional datasets and are yet not lifecycle enhanced. In contrast, the EAC-DRFC approach has addressed the shortcomings of previous approaches by improving the process of feature selection and adaptability, which resulted in improved accuracy and reliability.

5. Conclusion

The proposed EAC-DRFC approach addresses these limitations by integrating EAC Optimization for optimal feature selection and hyperparameter tuning, along with DRFC for adaptive learning. Data normalization is done using the Min-Max method in the preprocessing of the dataset to ensure that all attributes are scaled uniformly, and biases resulting from different magnitudes are eliminated. Additionally, PCA is performed on the data for attribute selection purposes, thereby retaining the important attributes. The proposed model achieves superior performance with 0.97 accuracy, 0.96 precision, 0.98 recall, and a 0.97 F1-score, demonstrating its effectiveness. The model's limits stemmed from the requirement for high-quality data and intricate calculations during optimization. Preparing the model was widely used, but it required a long time to train. Future concerns include edge computing methods, the usage of lightweight models, and the incorporation of current-time IoT data. The model could be improved by incorporating multi-objective optimization and DL.

Reference

1 Kasassbeh, S.A., Assolie, A.A., Qashmar, D., and Aljaberi, M., 2025. Sustainable Solutions: Advanced Analytics for Green Buildings Performance Using Machine Learning Models. Architecture, 13(6), pp.4508-4533. https://doi.org/10.13189/cea.2025.130628

2 Ige, A.B., Kupa, E. and Ilori, O., 2024. Best practices in cybersecurity for green building management systems: Protecting sustainable infrastructure from cyber threats. International Journal of Science and Research Archive, 12(1), pp.2960-2977. https://doi.org/10.30574/ijsra.2024.12.1.1185

3 AbdelHaffez, A.G., Kaloop, M.R., Eldessouki, M. and Issa, U.H., 2025. Identification and Assessment of Risk Factors in Green Building Projects: A Multi-Dimensional Approach for Sustainable Infrastructure. Sustainability, 17(22), p.10178. https://www.mdpi.com/2071-1050/17/22/10178#

4 Bungau, C.C., Bungau, T., Prada, I.F. and Prada, M.F., 2022. Green buildings as a necessity for sustainable environment development: dilemmas and challenges. Sustainability, 14(20), p.13121. https://www.mdpi.com/2071-1050/14/20/13121#

5 Herath, P. and Bai, X., 2024. Benefits and co-benefits of urban green infrastructure for sustainable cities: six current and emerging themes. Sustainability Science, 19(3), pp.1039-1063. https://doi.org/10.1007/s11625-024-01475-9

6 Mahmood, S., Misra, P., Sun, H., Luqman, A. and Papa, A., 2025. Sustainable infrastructure, energy projects, and economic growth: mediating role of sustainable supply chain management. Annals of operations research, 355(1), pp.1099-1130. https://doi.org/10.1007/s10479-023-05777-6

7 Mahmood, S., Sun, H., Iqbal, A., Alhussan, A.A. and El-kenawy, E.S.M., 2024. Green finance, sustainable infrastructure, and green technology innovation: pathways to achieving sustainable development goals in the belt and road initiative. Environmental Research Communications, 6(10), p.105036. https://doi.org/10.1088/2515-7620/ad898f

8 Liu, T., Chen, L., Yang, M., Sandanayake, M., Miao, P., Shi, Y. and Yap, P.S., 2022. Sustainability considerations of green buildings: a detailed overview on current advancements and future considerations. Sustainability, 14(21), p.14393. https://www.mdpi.com/2071-1050/14/21/14393#

9 Karimi, H., Adibhesami, M.A., Bazazzadeh, H. and Movafagh, S., 2023. Green buildings: Human-centered and energy efficiency optimization strategies. Energies, 16(9), p.3681. https://doi.org/10.3390/en16093681

10 Gil-Ozoudeh, I., Iwuanyanwu, O., Okwandu, A.C. and Ike, C.S., 2023. Sustainable urban design: The role of green buildings in shaping resilient cities. International Journal of Applied Research in Social Sciences, 5(10), pp.674-692. https://doi.org/10.51594/ijarss.v5i10.1481

11 Meena, C.S., Kumar, A., Jain, S., Rehman, A.U., Mishra, S., Sharma, N.K., Bajaj, M., Shafiq, M. and Eldin, E.T., 2022. Innovation in green building sector for sustainable future. Energies, 15(18), p.6631. https://www.mdpi.com/1996-1073/15/18/6631#

12 Yang, B., Lv, Z. and Wang, F., 2022. Digital twins for intelligent green buildings. Buildings, 12(6), p.856. https://doi.org/10.3390/buildings12060856

13 Bashir, M.T., Khan, A.B., Khan, M.M.H., Rasheed, K., Saad, S. and Farid, F., 2024. Evaluating the implementation of green building materials in the construction sector of developing nations. J. Hum. Earth Future, 5(3), pp.528-542. https://doi.org/10.28991/HEF-2024-05-03-015

14 Iwuanyanwu, O., Gil-Ozoudeh, I., Okwandu, A.C. and Ike, C.S., 2024. Cultural and social dimensions of green architecture: Designing for sustainability and community well-being. International Journal of Applied Research in Social Sciences, 6(8), pp.1951-1968. https://doi.org/10.51594/ijarss.v6i8.1477

15 Pragati, S., Shanthi Priya, R., Pradeepa, C. and Senthil, R., 2023. Simulation of the energy performance of a building with green roofs and green walls in a tropical climate. Sustainability, 15(3), p.2006. https://doi.org/10.3390/su15032006

16 Mahmood, S., Sun, H., El-Kenawy, E.S.M., Iqbal, A., Alharbi, A.H. and Khafaga, D.S., 2024. Integrating machine and deep learning technologies in green buildings for enhanced energy efficiency and environmental sustainability. Scientific Reports, 14(1), p.20331. https://doi.org/10.1038/s41598-024-70519-y