Proceedings of Machine Learning Research vol 245:1–7, 2024 2024 International Conference on Machine Learning and Intelligent Computing

Research on Green Design Optimization of Ethnic Minority Architecture in Guangxi Based on Machine Learning

Cong Lu* 502208074@QQ.COM
College of Architecture and Civil Engineering, Nanning University, Nanning, Guangxi, China
Nenglang Huang
College of Architecture and Civil Engineering, Nanning University, Nanning, Guangxi, China
Yang Wu
College of Architecture and Civil Engineering, Nanning University, Nanning, Guangxi, China

Editors: Nianyin Zeng and Ram Bilas Pachori

Abstract

Guangxi Zhuang Autonomous Region, as one of China's ethnic minority areas, possesses rich heritage resources in ethnic architecture, with dual objectives of cultural preservation and ecological sustainability. This paper explores the design principles of ethnic minority architecture in Guangxi, investigates the relationship between architecture and climate adaptability, and integrates them with digital fabrication technology. Utilizing parametric platforms and performance simulation tools, the study examines the climate adaptability of ethnic minority architecture in Guangxi. Through machine learning, models for lighting, thermal, and humidity environments specific to Guangxi's ethnic minority regions are developed. Optimization parameters for architectural design are proposed, and the reliability and accuracy of the models are demonstrated through training and testing, providing ecological design optimization strategies and references for future research on green architecture in ethnic minority areas.

Keywords: Ethnic Architecture; Green Design; Machine Learning; Guangxi Region

1. Introduction

With the rapid development of socio-economy and the acceleration of urbanization worldwide, the construction industry plays an increasingly crucial role. However, this growth brings forth issues such as excessive resource consumption and environmental pollution, highlighting the urgent need for sustainable solutions. In 2024, the Ministry of Housing and Urban-Rural Development issued a notice to expedite energy-saving and carbon reduction efforts in the construction sector, emphasizing the comprehensive implementation of green building standards for new constructions (B. et al., 2023). As a significant part of Southwest China, Guangxi Zhuang Autonomous Region boasts rich ethnic minority cultures and unique geographical environments, offering abundant resources for green architectural design. Nevertheless, the lack of systematic scientific methods and technological means poses challenges to green design and optimization in ethnic minority architecture in Guangxi.

Machine learning, as a powerful data analysis tool, has seen increasingly wide applications in the field of architecture in recent years. By employing machine learning algorithms to simulate and optimize building environments, it can effectively enhance energy utilization efficiency and indoor comfort in architectural design (Kreuzberger et al., 2023). This paper aims to utilize machine learning techniques to explore optimization strategies for green building design based on the characteristics and environmental conditions of ethnic minority architecture in Guangxi. The objective is to provide scientific support and technical guidance for the green development of architecture in the Guangxi region.

2. Overview of Machine Learning

2.1. Principles of Machine Learning

Machine learning is widely used in the fields of image recognition, speech recognition, natural language processing, and predictive analytics. In 1959, Arthur Semel first introduced the concept of "machine learning" and defined it as a field that enables computers to learn on their own without precise programming instructions. Machine learning algorithms can be categorized into three types: supervised learning, unsupervised learning, and reinforcement learning. Supervised learning involves training on labeled data, where the data is already labeled with correct outputs. The algorithm then uses this labeled data to learn how to predict outputs for new, unseen data. Unsupervised learning refers to training algorithms on unlabeled data, where the algorithm searches for patterns in the data and groups similar data points together. Reinforcement learning involves learning through trial and error, receiving feedback in the form of rewards or penalties, and learning to take actions that maximize rewards over time. Machine learning algorithms capture nonlinear relationships among various factors that influence predictive outcomes.

2.2. Machine Learning for Green Design

In building energy management, machine learning algorithms are extensively employed for the analysis, prediction, and optimization control of building energy consumption data. Utilizing random forest algorithms for analyzing and predicting building energy consumption data enables precise forecasting and energy management (Kapp et al., 2023). Furthermore, deep learning algorithms facilitate intelligent optimization control of building energy systems, allowing for adaptive adjustment and optimized operation of building energy systems.

In architectural design, the support vector machine algorithm predicts and analyses the performance data of building materials, providing a scientific basis for the selection and application of green building materials. In addition, the genetic algorithm carries out automatic generation and optimal selection of building appearance design, which can improve the efficiency of building design and design quality.

In the domain of building environment monitoring and intelligent control, machine learning algorithms are applied to the monitoring of indoor environmental data and the design of intelligent control systems. Deep learning algorithms classify and recognize indoor environmental data, enabling intelligent monitoring and control of building interiors. Reinforcement learning algorithms optimize building intelligent control systems to ensure intelligent adjustment and optimized operation of building equipment.

The application of machine learning algorithms in the field of green building encompasses various aspects including building energy management, material selection, and design optimization, providing new technical support and solutions for the development and promotion of green buildings. However, current research still faces challenges such as low data quality and the need to improve algorithm stability, requiring further exploration and investigation.

3. Guangxi Minority Architectural Features

3.1. Construction Form

Guangxi is located in the subtropical monsoon climate zone, with complex and varied topography, undulating mountain ranges and rivers, forming a rich natural scenery and diversified regional characteristics. Minority architecture is often integrated with the local natural environment, making full use of local natural resources and topography to build houses according to local conditions, forming a unique style of patio-style residential architecture.

The ”one bright, two dark” layout with three openings is the most basic residential pattern of traditional dwellings in northern Guangxi. To address the issue of limited space in the ”one bright, two dark” layout and the hot and humid climate of Guangxi, second floors are often added to the left and right side rooms to prevent moisture. Traditional dwellings feature axial symmetry in their architectural layout, utilizing mortise and tenon jointed timber frames to support the internal space. This construction method offers good thermal performance, effectively reducing heat transfer from outdoors to indoors, and minimizing heat loss indoors during winter (Zhou et al., 2024). To address drainage issues, roofs are designed in a herringbone shape with a significant slope, and lightweight breathable materials such as small blue tiles are used to alleviate indoor humidity problems.

In traditional dwellings in Guangxi, the doors and windows facing the courtyard are often adorned with hollow lattice patterns, serving not only to effectively divide spaces but also seamlessly connect with open halls and courtyards to enhance indoor ventilation. Main doors, outer doors, and room doors are simple solid panel doors. Under feudal systems, traditional dwelling walls are typically high, hindering ventilation. To address this, several small square latticed windows are installed at the top of high walls, with a height of only 30-50cm. These windows feature various patterns, meeting ventilation and lighting needs.

3.2. Construction Material

The building materials used in Guangxi’s ethnic minority architecture are diverse and unique, including wood, bamboo, stone, adobe and reeds. The region’s abundant forest resources provide ample sources of wood for ethnic minorities. Wood is commonly used for constructing building structures such as beams and columns, as well as for making furniture and decorations, boasting natural beauty and environmental friendliness. Bamboo also holds significance in Guangxi’s ethnic minority architecture, prized for its lightweight, strength, and durability, commonly utilized for structural and decorative purposes. In construction, bamboo is often used for roofing, railings, and flooring, as well as for crafting woven goods and handicrafts such as bamboo furniture and baskets. Adobe, made from soil, sand, and hemp ropes, is simple to produce, cost-effective, and effectively regulates indoor temperature and humidity while offering good insulation properties. Additionally, stone is commonly used for building foundations and wall structures, known for its durability, enhancing stability and longevity of buildings.

4. Optimization Strategy for Green Design of Ethnic Minority Architecture in Guangxi Based on Machine Learning

The supervised learning (Hu and Szymczak, 2023) idea in machine learning is used to construct the Guangxi minority building characteristics, and the thermal environment and light environment and humidity conditions in the climate environment are incorporated into the model for research and calculation.The core of the predictive model involves utilizing a simulation database for lighting, thermal, and humidity environments, trained with the XGBoost model. Subsequently, using cross-validation and hyperparameter tuning methods within a defined parameter space, we optimized the hyperparameters to generate a high-precision predictive model for the thermal, lighting, and humidity environments of ethnic minority architecture in Guangxi (Mucha, 2020).

Extreme gradient boosting (XGBoost) is an upgraded version of the gradient boosting tree algorithm. It is an ensemble algorithm based on decision trees that iteratively reduces the residuals from the previous iteration by constructing base learners. The final prediction output is obtained by summing the results of multiple base learners (Ozaki et al., 2020)(as shown in Figure 1). XGBoost achieves a balance between model accuracy and complexity through loss functions and regularization terms, thus improving model efficiency while ensuring performance. Performance evaluation metrics for predictive models include the coefficient of determination (R2), root mean square error (RMSE), and mean absolute error (MAE) (Zhang et al., 2016). R2 assesses the feasibility and linearity of the model, MAE calculates the average absolute error between predicted and actual values, and RMSE evaluates the accuracy of regression model predictions.

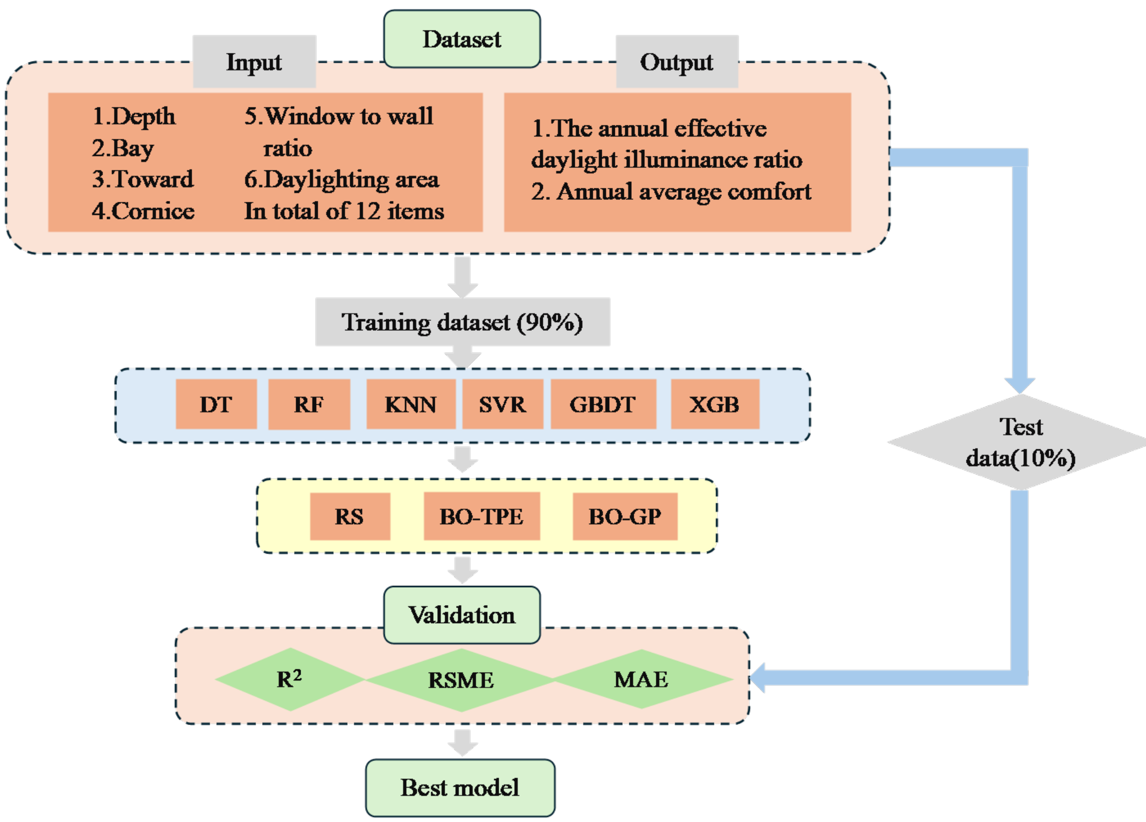

Figure 1: Flowchart of machine learning

To enhance the precision of model learning, preprocessing is essential before employing machine learning methods for data analysis. By performing tasks such as cleaning, transformation, integration, and reduction on raw data, data quality is improved, leading to enhanced accuracy of the model. The purpose of data preprocessing is to prepare the data for better handling and analysis by machine learning algorithms, thus achieving improved predictive performance (D'Agostino et al., 2023). However, this dataset is incompatible with machine learning algorithms in the Scikit-learn library because machine learning algorithms require all values in the dataset to be numeric and meaningful. To meet this requirement, the dataset needs to undergo cleaning operations (Krarti, 2003). The original dataset contains a total of 521 data samples, but after cleaning, only 463 valid data samples remain.

In this study, the 463 data samples will be divided into training and testing sets using a 9:1 ratio. The training set, comprising 417 data samples, will be used to train the predictive model, while the testing set, containing 46 data samples, will be used to evaluate the model's predictive performance. During training, 10-fold cross-validation will be employed to enhance the model's generalization ability and improve its performance, as shown in the Figure 2.

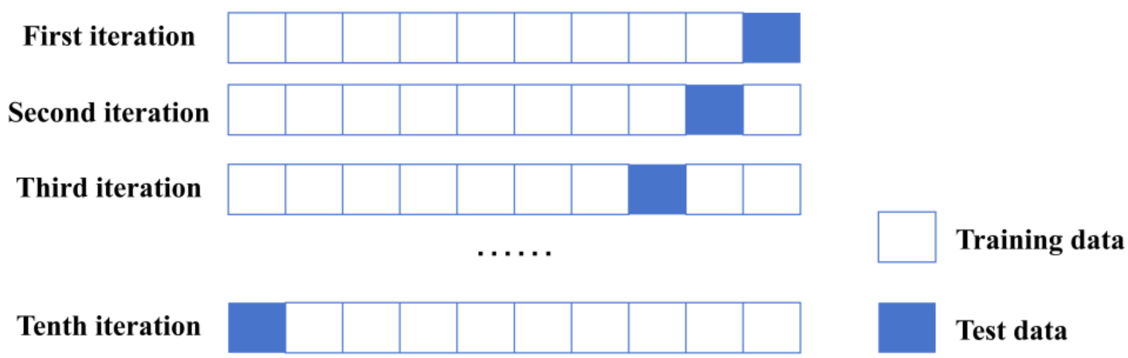

Figure 2: Cross-validation diagram

The Bayesian approach based on Gaussian process regression exhibits good performance in complex XGBoost algorithms. The specific hyperparameter metrics obtained are shown in Table 1.

Table 1: Combination of hyperparameters for XGBoost method configuration

Algorithm	Hyperparameter	Random search
XGBoost	estimator__gamma	0.005
	estimator__learning_rate	0.654
	estimator__max_depth	12
	estimator__n_estimators	63
	estimator__reg_lambda	0.987
	estimator__subsample	0.636

The difference values and test values of light, heat and humidity environments predicted by different algorithmic models for Guangxi ethnic minority buildings are shown in Table 2.

The results from Table 2 indicate that the model's predictive accuracy has met expectations, with parameter bias rates all below 10%. The reliability of the model has been verified, which can serve as an algorithmic foundation for optimizing more related parameters in the green design of ethnic minority architecture in Guangxi in the future.

Table 2: Predictive model deviation rate and test value

	R^2			RMSE	Deviation rate
	Training	Testing	Cross-validation		
Light environment	0.9776	0.9655	0.9467	0.1120	3%
Thermal environment	0.9132	0.9531	0.9211	0.0398	5%
Humid environment	0.8948	0.8772	0.8991	0.4055	6%

The above study found that adopt a collaborative design approach in future projects to enhance the green performance of ethnic minority architecture in Guangxi. Utilizing collaborative design, building design schemes can be optimized based on indoor environmental parameters and energy consumption requirements specified by relevant regulations. The renovation and redesign process should prioritize climate adaptability, aiming to reduce heating and cooling energy consumption by leveraging passive architectural design techniques such as natural lighting, natural ventilation, and thermal insulation in the building envelope. This approach aims to elevate the green standards and ecological sustainability of ethnic minority architecture in Guangxi.

5. Conclusions

In this study, a machine learning predictive model for the annual effective daylighting illuminance ratio of ethnic minority architecture in Guangxi was established using 463 optimized solution data exported from the multi-objective generation module. Bayesian optimization was employed to adjust the hyperparameters of the machine learning model, leading to the following conclusions:

(1) Based on Bayesian optimisation hyper-parameter configuration method, the performance of multi-regression prediction model for annual effective lighting illuminance share and annual average thermal comfort of Guangxi ethnic minority buildings is effectively improved, and the efficiency is higher than other hyper-parameter optimisation methods.

(2) The machine learning prediction model for the annual effective light illuminance percentage of traditional residential houses proposed in this paper can be used in the design of Guangxi ethnic minority buildings or the renovation of residential houses, which facilitates the subsequent determination of the design parameters and contributes to the improvement of the performance effect of the design scheme.

This study aligns with the current trend of energy-saving and green building development, addressing prominent issues in the green performance of ethnic minority architecture in Guangxi. By developing a performance-driven parameter optimization design method and employing computer simulation optimization tools combined with machine learning methods, design strategies were obtained.

Acknowledgments

The authors acknowledge the financial support from Young and middle-aged teachers of Guangxi Universities in 2022 research basic capacity Improvement Project: Research on parametric design and production of ethnic minority buildings in Guangxi based on three-dimensional reconstruction (No. 2023KY194).

References

Sangmesh B., Nagraj Patil, Krishna Kumar Jaiswal, T.P. Gowrishankar, K. Karthik Selvakumar, M.S. Jyothi, R. Jyothilakshmi, and Santosh Kumar. Development of sustainable alternative materials for the construction of green buildings using agricultural residues: A review. *Construction and Building Materials*, 368:130457, 2023. doi: 10.1016/j.conbuildmat.2023.130457.

Diana D'Agostino, Federico Minelli, and Francesco Minichiello. New genetic algorithm-based workflow for multi-objective optimization of net zero energy buildings integrating robustness assessment. *Energy and Buildings*, 2023.

Jianchang Hu and Silke Szymczak. A review on longitudinal data analysis with random forest. *Briefings in Bioinformatics*, 24(2):bbad002, 01 2023. doi: 10.1093/bib/bbad002.

Sean Kapp, Jun-Ki Choi, and Taehoon Hong. Predicting industrial building energy consumption with statistical and machine-learning models informed by physical system parameters. *Renewable and Sustainable Energy Reviews*, 172:113045, 2023. doi: 10.1016/j.rser.2022.113045.

Moncef Krarti. An Overview of Artificial Intelligence-Based Methods for Building Energy Systems . *Journal of Solar Energy Engineering*, 125(3):331–342, 08 2003. doi: 10.1115/1.1592186.

Dominik Kreuzberger, Niklas Kühl, and Sebastian Hirschl. Machine learning operations (mlops): Overview, definition, and architecture. *IEEE Access*, PP:1–1, 01 2023. doi: 10.1109/ACCESS.2023.3262138.

Waldemar Mucha. Comparison of machine learning algorithms for structure state prediction in operational load monitoring. *Sensors*, 20(24), 2020. doi: 10.3390/s20247087.

Yoshihiko Ozaki, Yuki Tanigaki, Shuhei Watanabe, and Masaki Onishi. Multiobjective tree-structured parzen estimator for computationally expensive optimization problems. In *Proceedings of the 2020 Genetic and Evolutionary Computation Conference*, GECCO '20, page 533–541, New York, NY, USA, 2020. Association for Computing Machinery. doi: 10.1145/3377930.3389817.

Longwei Zhang, Lingling Zhang, and Yuetao Wang. Shape optimization of free-form buildings based on solar radiation gain and space efficiency using a multi-objective genetic algorithm in the severe cold zones of china. *Solar Energy*, 132:38–50, 2016. doi: 10.1016/j.solener.2016.02.053.

Yuchen Zhou, Mu Liu, Guanhong Xie, and Chunqing Liu. Landscape ecology analysis of traditional villages: A case study of ganjiang river basin. *Applied Sciences*, 14(2), 2024. doi: 10.3390/app14020929.