GREEN BUILDING PROJECT MANAGEMENT APPROACH IN CONSTRUCTION PHASES

Aachal Bhaurao Kore[1, a *] and Dr. Amitkumar B. Ranit[2, b]

[1] Construction Engineering and Management, Prof. Ram Meghe College of Engineering and Management, India

[2] Associate Professor, Department of Civil Engineering, Prof. Ram Meghe College of Engineering and Management, India

Abstract

The green building market is an essential component of global sustainability efforts, aiming to reduce energy consumption, carbon emissions, and resource depletion while promoting eco-friendly construction practices. However, achieving these goals is often hindered by challenges such as complex project delivery systems, cost management issues, and difficulties in meeting certification standards. This study explores the application of machine learning, specifically Artificial Neural Networks (ANNs), to address these challenges and enhance the efficiency and effectiveness of green building initiatives. Using a dataset comprising key performance indicators (KPIs) such as revenue, rank, staff accreditation, and market diversity, this research develops and trains an ANN model to classify green building companies into performance clusters. The model leverages features like average green revenue and percentage of green revenue to identify high-performing, mid-performing, and low-performing companies. Metrics such as accuracy (90.94%), F1 score (0.05), and performance loss (3.35) are used to evaluate the model, revealing its strengths and limitations in handling imbalanced datasets. This paper demonstrates the potential of machine learning to provide actionable insights for stakeholders, optimize resource allocation, and improve compliance with green certification standards. By integrating data-driven decision-making into the green building lifecycle, this study highlights how machine learning can drive innovation, enhance sustainability, and contribute to global climate goals. Future work focuses on expanding datasets, improving model precision, and integrating real-time data analytics for continuous performance optimization.

Keywords: Green Building Market, Artificial Neural Networks (ANN), Machine Learning, Sustainability, Energy Efficiency

Introduction

The construction industry plays a pivotal role in global economic development but is also one of the largest contributors to environmental degradation. Buildings account for approximately 37% of global carbon dioxide (CO_2) emissions and 34% of energy demand, making it imperative to adopt sustainable practices to mitigate these impacts. Green buildings, designed to optimize energy efficiency, reduce resource consumption, and minimize environmental footprints, have emerged as a critical solution to address these challenges. However, the implementation and widespread adoption of green building practices remain complex due to financial, technical, and regulatory barriers.

In this context, machine learning (ML) has emerged as a transformative technology capable of addressing these complexities. By analyzing vast datasets and uncovering patterns, ML enables stakeholders in the green building market to make informed decisions, optimize resources, and improve compliance with sustainability standards. Artificial Neural Networks (ANNs), a subset of ML, are particularly effective in modeling non-linear relationships and classifying data, making them a valuable tool for advancing green building initiatives.

This study focuses on leveraging ANNs to classify green building companies based on performance metrics, including revenue, rank, staff accreditation, and market diversity. Using these classifications, stakeholders can identify high-performing companies as benchmarks, understand gaps in mid- and low-performing companies, and develop strategies to enhance overall market efficiency. By integrating data-driven approaches with sustainable development goals, the project seeks to address critical challenges in project delivery, cost management, and energy efficiency

Literature

Over the past few decades, scholars have increasingly focused on the research on green building [1], and have issued an increasing number of papers [2]. This may make it tough to grasp the research focus and status quo from thousands of papers, posing a major risk of neglecting essential questions and areas for research and practice improvement [3]. In order to solve this problem, it is necessary to analyze this field by utilizing scientometric software [4]. A literature review is considered to be an effective way to deeply understand the field of research [5]. By systematically combing the existing research, we can figure out the current research situation and development trend of the field, thus providing a direction for future research [5]. It should be pointed out that the development of knowledge is a dynamic process. As scientific literature is constantly updated, we may not have enough time or effort to track it only by relying on non-visualization technology.

With the development of science and technology, many visualization tools have emerged in recent years, such as VOSviewer, CoPalRed, Bibexcel, Sci2, VantagePoint, and CiteSpace. All of these tools support document co-citation analysis and keyword co-occurrence analysis, which can help us conduct quantitative and objective analysis of the relevant fields, and reveal the quantitative relations among various studies. For example, Li et al. carried out document co-citation analysis and cluster analysis on the relevant literature from 2004 to 2015 via CiteSpace, and quantitatively proposed the building information modeling (BIM) knowledge graph [6]. On the basis of the Web of Science (WoS) databases, Jiang et al. figured out the research emphasis and development trend of urban planning for climate change from 1990 to 2016 through cluster analysis and knowledge evolution analysis using CiteSpace [7]. Zhao et al. analyzed the characteristics and trends of the new energy vehicle reliability based on literature from 1998 to 2017 using CiteSpace [8]. Chen et al. analyzed 3875 articles related to regenerative medicine from 2000 to 2011 using CiteSpace, finding emerging trends in this area [9].

Problem Statement

Despite the growing demand for green buildings, significant barriers persist, including:

1. High initial costs and long payback periods.
2. Technical challenges, such as limited availability of sustainable materials and skilled professionals.
3. Regulatory hurdles, including inconsistent policies and weak enforcement mechanisms.

These issues highlight the need for innovative, data-driven solutions to streamline green building operations and improve market outcomes.

Research Objectives

This study aims to:

1. Analyze key performance indicators (KPIs) in the green building market using a structured dataset.
2. Develop and train an ANN model to classify companies into performance clusters.
3. Evaluate the model's accuracy, F1 score, and performance loss to assess its effectiveness.
4. Provide actionable insights to optimize resource allocation, enhance sustainability practices, and improve compliance with green certification standards.

Significance of the Study

The integration of machine learning into the green building market has far-reaching implications:

1. **Economic Benefits:** Improved cost management and resource allocation.
2. **Environmental Impact:** Enhanced energy efficiency and reduced carbon emissions.
3. **Policy Support:** Data-driven recommendations for designing targeted incentives and regulatory frameworks.

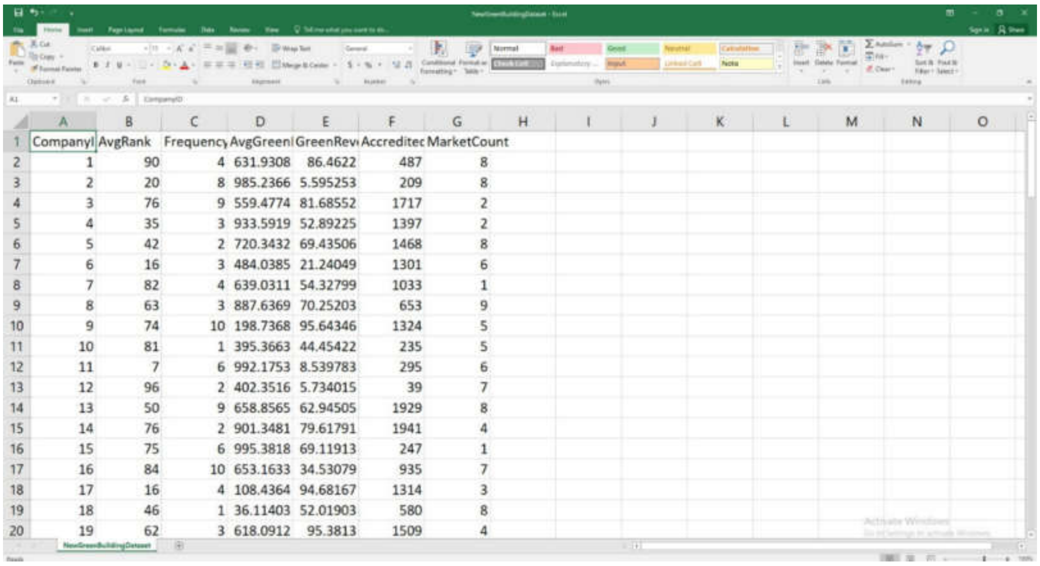
This study contributes to the growing body of research on sustainable development by demonstrating the potential of machine learning to revolutionize green building practices. The results of this research will provide a framework for stakeholders to achieve both economic and environmental goals, driving progress toward a more sustainable future.

Dataset

The dataset is designed to analyze and classify green building companies based on their performance metrics. Each feature provides critical insights into various dimensions of company performance, including financial success, operational capacity, market engagement, and technical expertise.

Application in the Project

- **Performance Clustering**: The features are used as input to an Artificial Neural Network (ANN) model to classify companies into high-performing, mid-performing, and low-performing clusters.
- **Feature Significance**: Attributes like AvgGreenRevenue, GreenRevenuePercentage, and AccreditedStaff are key indicators for identifying leaders in the green building market.
- **Data Analysis**: The dataset enables exploratory analysis to uncover patterns, such as the correlation between revenue and staff accreditation, or the relationship between rank and frequency of participation.
- **Insights for Stakeholders**: Helps stakeholders make data-driven decisions, such as identifying areas for improvement or benchmarking against high-performing companies.



	A	B	C	D	E	F	G
1	CompanyI	AvgRank	Frequency	AvgGreenI	GreenRev	Accreditec	MarketCount
2	1	90	4	631.9308	86.4622	487	8
3	2	20	8	985.2366	5.595253	209	8
4	3	76	9	559.4774	81.68552	1717	2
5	4	35	3	933.5919	52.89225	1397	2
6	5	42	2	720.3432	69.43506	1468	8
7	6	16	3	484.0385	21.24049	1301	6
8	7	82	4	639.0311	54.32799	1033	1
9	8	63	3	887.6369	70.25203	653	9
10	9	74	10	198.7368	95.64346	1324	5
11	10	81	1	395.3663	44.45422	235	5
12	11	7	6	992.1753	8.539783	295	6
13	12	96	2	402.3516	5.734015	39	7
14	13	50	9	658.8565	62.94505	1929	8
15	14	76	2	901.3481	79.61791	1941	4
16	15	75	6	995.3818	69.11913	247	1
17	16	84	10	653.1633	34.53079	935	7
18	17	16	4	108.4364	94.68167	1314	3
19	18	46	1	36.11403	52.01903	580	8
20	19	62	3	618.0912	95.3813	1509	4

Figure Dataset of excel sheet

Results

Artificial neural network

Graph show actual and Predicted Market Demand

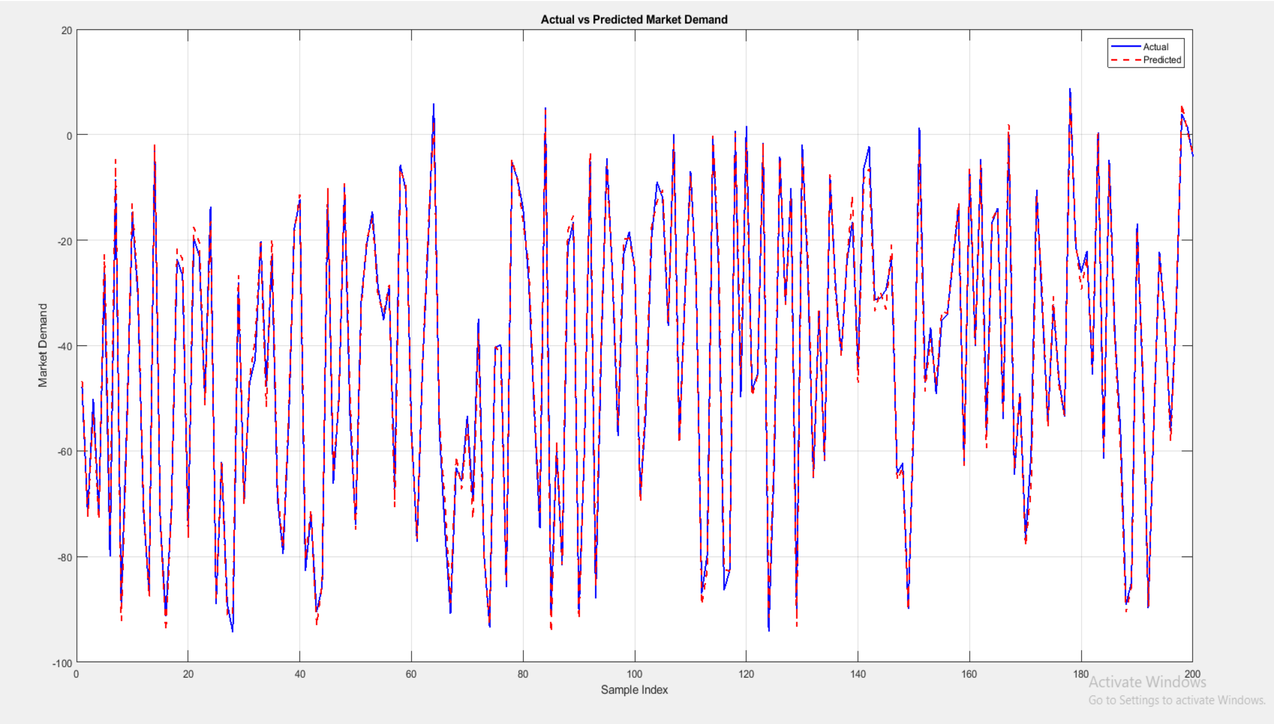


Figure. Graph show actual and Predicted Market Demand

The program generates a plot that visually compares the actual market demand values with the predicted values derived from the neural network. The plot employs distinct colors for actual and predicted data points, enabling quick visual assessment of model accuracy and the quality of the predictions made by the neural network.

Performance Evaluation

The performance of a machine learning model is evaluated using several key metrics that provide insights into its predictive accuracy, precision, recall, and overall effectiveness. In this case, we have three important performance metrics: **Loss (Performance)**, **Accuracy**, and **F1 Score**. These metrics are used to assess how well the trained model generalizes to unseen data.

1. Performance (Loss): 3.35

Performance, also referred to as loss in machine learning, quantifies how well the model's predictions match the true values. A lower value indicates better performance, as it means that the model's predictions are closer to the actual outcomes.

- **Interpretation**: In this case, the loss value of **3.35** represents the discrepancy between the predicted and actual values. While this number is somewhat high, it should be evaluated in the context of the specific task, dataset, and model used. Loss values are typically used to guide model optimization through training by minimizing this value through techniques like gradient descent.
- **Context**: A loss of 3.35 could indicate that the model is still learning, and further tuning or additional training epochs might be required to lower the loss. However, this value should be examined along with the accuracy and F1 score to understand the overall quality of the model.

2. Accuracy: 90.94%

Accuracy is one of the most commonly used metrics for classification models. It measures the percentage of correct predictions (both true positives and true negatives) out of all predictions made.

- **Formula**:

$$\text{Accuracy} = \frac{\text{Number of Correct Predictions}}{\text{Total Number of Predictions}} \times 100$$

- **Interpretation**: An accuracy of **90.94%** suggests that the model correctly predicted the outcomes for nearly 91% of the test dataset. This is considered a strong result, as it implies that the model is performing well at distinguishing between classes.

- **Consideration**: While a high accuracy is generally desirable, it is important to also consider other metrics, especially when dealing with imbalanced datasets or multiple classes. High accuracy can sometimes be misleading if the model is biased towards predicting the majority class, which is why metrics like the **F1 Score** are also important.

3. F1 Score: 0.05

The **F1 score** is a more nuanced metric that combines **precision** and **recall** into a single score. It is especially useful when dealing with imbalanced datasets, where one class might dominate the others. The F1 score is calculated as the harmonic mean of precision and recall:

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

- **Precision** refers to the number of true positives divided by the total number of predicted positives.
- **Recall** refers to the number of true positives divided by the total number of actual positives.
- **Interpretation**: An F1 score of **0.05** is very low, indicating that while the model may have a high accuracy, it is struggling with precision and recall, particularly for certain classes. This could be a sign of class imbalance, where the model might be biased toward the majority class and is failing to correctly predict the minority class.
- **Why It's Important**: In tasks where the classes are imbalanced, relying solely on accuracy can be misleading. A low F1 score, as seen here, indicates that the model has poor performance on at least one of the classes, despite having a high overall accuracy. The F1 score is a better reflection of a model's ability to correctly classify both the minority and majority classes.

Conclusion

The integration of machine learning, particularly Artificial Neural Networks (ANN), into the green building market represents a transformative step toward addressing critical challenges in sustainability, energy efficiency, and cost optimization. This project demonstrates the potential of data-driven solutions in revolutionizing how green building companies manage their operations, achieve certifications, and enhance their market performance.

This project underscores the importance of machine learning in driving innovation within the green building market. By enabling data-driven decision-making, enhancing resource optimization, and streamlining operations, ANN-based solutions provide a pathway toward more sustainable and efficient building practices. As the green building market continues to grow, such technologies will play a pivotal role in shaping its future and contributing to global sustainability efforts

References

[1] Darko, A.; Chan, A.P.C. Critical analysis of green building research trend in construction journals. Habitat Int. **2016**, 57, 53–63.

[2] Ulubeyli, S.; Kazanci, O. Holistic sustainability assessment of green building industry in Turkey. J. Clean. Prod. **2018**, 202, 197–212.

[3] Darko, A.; Chan, A.P.C.; Huo, X.S.; Owusu-Manu, D. A scientometric analysis and visualization of global green building research. Build. Environ. **2019**, 149, 501–511.

[4] Chen, C.M. CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. J. Am. Soc. Inf. Sci. Tec. **2006**, 57, 359–377.

[5] Zuo, J.; Zhao, Z.Y. Green building research–current status and future agenda: A review. Renew. Sustain. Energy Rev. **2014**, 30, 271–281.

[6] Li, X.;Wu, P.; Shen, G.Q.P.;Wang, X.Y.; Teng, Y. Mapping the knowledge domains of building information modeling (BIM): A bibliometric approach. Autom. Constr. **2017**, 84, 195–206.

[7] Jiang, Y.F.; Hou, L.Y.; Shi, T.M.; Gui, Q.C.Areview of urban planning research for climate change. Sustainability **2017**, 9, 2224.

[8] Zhao, X.;Wang, S.;Wang, X. Characteristics and trends of research on new energy vehicle reliability based on the web of science. Sustainability **2018**, 10, 3560.

[9] Chen, C.M.; Hu, Z.G.; Liu, S.B.; Tseng, H. Emerging trends in regenerative medicine: A scientometric analysis in CiteSpace. Expert Opin. Biol. Ther. **2012**, 12, 593–608.

[10] Schneider, J.W.Mapping scientific frontiers: The quest for knowledge visualization. J. Am. Soc. Inf. Sci. Technol. **2004**, 55, 363–365.