SUPTM 2026: 3rd Conference on Future Challenges in Sustainable Urban Planning & Territorial Management

Universidad Politécnica de Cartagena sciforum

Predicting Lobby Location as an Indicator of Occupant Comfort: A Machine Learning (ML) Approach

Md. Mashnur Hossain Pobon[1], Md. Faisal Rahman Dhrubo[1,*], Maisha Mahmud[2] and Ramisa Anjum[2]

[1] Department of Civil Engineering, IUBAT—International University of Business Agriculture & Technology, Dhaka, Bangladesh; 22206078@iubat.edu; 22206057@iubat.edu
[2] Department of Agriculture, IUBAT—International University of Business Agriculture & Technology, Dhaka, Bangladesh; 23109073@iubat.edu; 23309167@iubat.edu
* Correspondence: 22206057@iubat.edu; Tel.: +880 1312360691

Abstract: This study investigates occupant comfort and spatial preference in hospital lobbies of Dhaka, Bangladesh, addressing the gap in data-driven analysis of hospital lobby spaces. Environmental, spatial, and behavioral data were collected and analyzed using three Machine Learning (ML) models: Decision tree (DT), Random Forest (RF), and Extreme Gradient Boosting (XGBoost). RF achieved the highest accuracy of 0.96 with RMSE of 0.65, identifying key factors affecting the lobby satisfaction. SHAP analysis showed that spatial attributes, especially Floor_Level and Window_State, are dominant over environmental conditions. This study shows the potential of ML to support Evidence-Based (EB) user-centered design in urban hospitals.

Keywords: Hospital lobby; Evidence-based guidance; Occupant comfort; Machine learning

1. Introduction

Artificially ventilated hospital lobbies have a significant influence on visitors' first impressions and psychological comfort. They are crucial transitional areas that connect the urban environment with healthcare facilities. In order to lower stress and ensure user wellbeing, elements like thermal comfort, illumination, and air quality are essential. However, high density and low Indoor Environmental Quality (IEQ) are common problems for hospitals in rapidly urbanizing places like Dhaka. Despite the significance of these regions, there is very little study relating occupant behavior to physical circumstances.

By using machine learning models - Decision Tree, Random Forest, and XGBoost - to predict lobby location preferences as a measure of occupant well-being, this study closes that gap. The aim of the research is to propose flexible, occupant-focused design techniques by examining the major variables affecting spatial comfort. These results provide useful recommendations for enhancing healthcare settings, ensuring that lobbies effectively satisfy user needs in the face of challenging urban conditions.

2. Literature Review

Hospital lobby design has an important effect on the wellbeing of its patrons. Ulrich's groundbreaking research showed that regenerative environments could reduce stress and improve healing [1]. The importance of lighting, air quality, and spatial design for psychological happiness in these transitional environments has been confirmed by additional study [2]. Zaman emphasizes that hospitals in Dhaka frequently have poor ventilation and high occupancy [4], but Frontzak and Wargocki identified thermal and air quality as important satisfaction drivers [3]. Furthermore, orientation and psychological security are impacted by spatial arrangement [5].

When taken as a whole, these features influence user comfort. While machine learning techniques such as DT, RF, and XGBoost provide enhanced prediction accuracy compared to conventional analysis [6], their use in this context is limited. There is still a big gap in creating data-driven models for hospital lobbies like Dhaka, despite the availability of interpretability tools like SHAP [7]. This underscores the necessity for tailored, evidence-based design solutions.

3. Methodology

This study utilizes a systematic, data driven methodology to forecast hospital lobby location preferences as a measure of occupant satisfaction, integrating environmental, spatial and behavioral variables via ML approaches. The overall workflow, represented in figure 1 has been broken down into three phases: data collection, model development and evaluation. The study was performed in hospital lobby of a multi storied healthcare

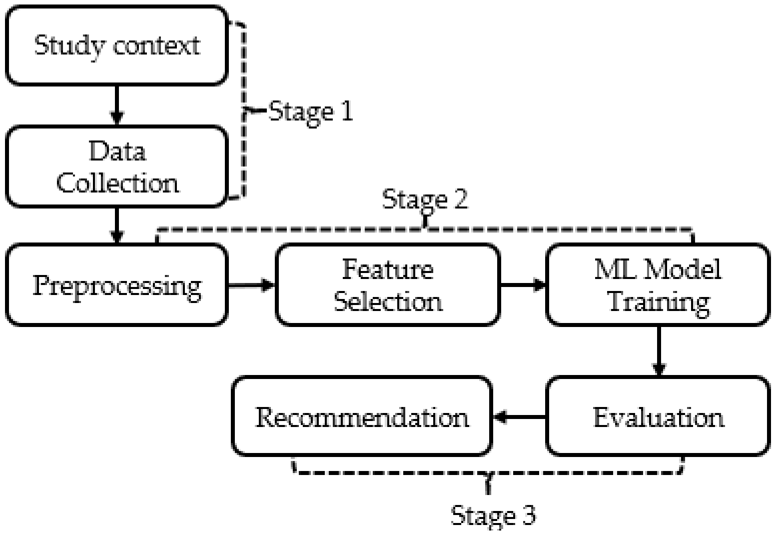

Figure 1: Methodology Flowchart

institution in Dhaka, Bangladesh, noted for its dynamic transitional space marked by high users' density, fluctuating environment, as well as distinct spatial arrangements. Data were gathered throughout the rainy season to ensure climatic stability, concentrating on operational hours from 10:00 to 17:00. A total of 400 reliable samples were collected utilizing organized questionaries combined with smart environmental monitoring instruments. The dataset included 24 variables encompassing demographic factors (i.e., gender, age, profession), environmental parameters (i.e., temperature, humidity, CO2 level and lighting level) spatial attributes (i.e., window status, air conditioning, seating type and floor level) and behavioral elements (i.e., waiting time, sitting position, and perceived crowding). Following data collection, preparation was conducted to assure data integrity, encompassing cleaning, normalization and label encoding of categorical features.

Outliers were analyzed and addressed without the implementation of oversampling or synthetic balancing methods, therefore maintaining the datasets inherent distribution. Three controlled ML models DT, RF and XGBoost were used due to their simplicity, reliability, and the possibility to answer the problem of the interpretation of complex relationships. The model was developed in Python using scikit learn and XGBoost libraries. The dataset was divided into training and test set using train-test spilt method to evaluate models' performance. Cross validation techniques such as K-Fold and GridSearchCV were applied to optimize RF model parameters and ensure robust, unbiased performance result. To measure the model's performance Accuracy, Mean Squared Error (MSE), Root Mean Squared Error (RMSE) was used. Additionally, SHAP analysis was performed to elucidate feature contributions and ascertain the most significant predictors. The comprehensive techniques yielded valid predicted outcomes and important insights into the impact of environmental and spatial conditions on users' comfort, consequently facilitating evidence based, occupant centered design plans for hospital lobbies in Dhaka, Bangladesh.

4. Result and Discussion

Three supervised ML model-DT, RF and XGBoost-were applied to predict hospital lobby location preference as an indicator of occupant comfort. Each model was trained on 80% of the dataset and tested on 20% using an identical set of environmental, spatial and behavioral features. Model performance was using accuracy, mean squared error (MSE), and root mean squared error (RMSE), with SHAP employed for model interpretation. As shown in Table 1, the RF model achieved the highest performance, with an accuracy of 0.96, MSE of 0.42, and RMSE of 0.65, demonstrating its effectiveness in capturing relationships among environmental, demographic, and spatial variables. XGBoost followed closely with an accuracy of 0.95, MSE of 0.54, and RMSE of 0.73, while the Decision Tree model showed comparatively lower performance (accuracy 0.91, MSE 0.76, RMSE 0.87), indicating higher prediction error.

Table 1. ML Model Performance

Parameters	ML Model	Accuracy	MSE	RMSE
Lobby Location	RF	0.96	0.42	0.65
	XGBoost	0.95	0.54	0.73
	DT	0.91	0.76	0.87

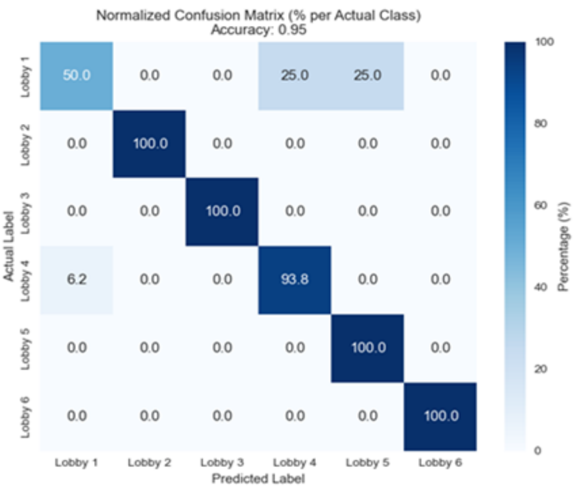

Figure 2: Confusion Matrix

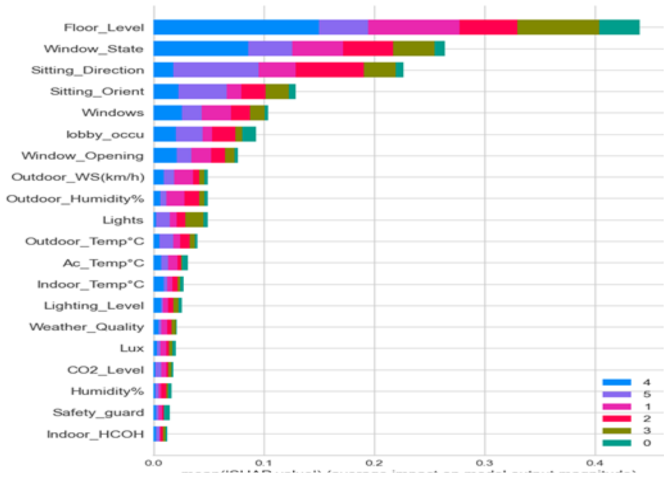

Figure 3: SHAP Analysis

The normalized confusion matrix in Figure 2 shows the classification performance across six lobby types. Each cell represents the percentage of samples from an actual lobby class (rows) predicted as a given lobby class (columns). Most predictions lie along the diagonal, indicating correct classifications. Misclassifications occur mainly for lobby 1 where 25% of the samples are confused with lobby 4 and another 25% of the samples with lobby 5. All other lobbies have near-perfect predictions. The overall model accuracy is 0.95, showing high reliability. The SHAP summary in Figure 3 plot demonstrates the relative importance of each feature in predicting lobby classifications. The horizontal axis shows the mean SHAP value, indicating how much each variable contributes, on average, to the model's output. Feature at top have the greatest influence, while those lower down have smaller impacts. The colored segments correspond to different lobby classes (0-5), showing how each feature affects various types differently. Floor_Level stands out as the most influential factor, meaning the floor position strongly influencing the model's decision.

The analysis showed that the building's layout plays the most significant role in distinguishing lobby types. Window conditions come next in importance, followed by how seats are arranged and oriented, reinforcing the importance of spatial design. Window placement and crowd density also shape how each lobby perceived. Environmental factors like wind, humidity, and lighting have a moderate impact, while CO_2 level and other air quality variables exhibit minimal influence. Overall, the results suggests that people respond more to design features than to short-term environmental changes. The SHAP plot highlights layout, orientation, and window opening features as the strongest predictors of lobby location.

4. Conclusion

This study investigated occupant experience in hospital lobbies using an integrated machine learning (ML) framework combining environmental, spatial, and behavioral parameters. Conducted in a 15-storied hospital in Dhaka, Bangladesh, three supervised models—Decision Tree (DT), Random Forest (RF), and XGBoost—were applied to predict lobby preference as an indicator of comfort. These models were selected for their ability to handle mixed real-time datasets, capture non-linear relationships, and provide interpretability. Among them, the Random Forest (RF) model achieved the highest performance (accuracy = 0.96, RMSE = 0.65), addressing limitations observed in the literature. SHAP analysis indicated that architectural and spatial attributes were more influential than environmental conditions, with Floor_Level, Window_State, Sitting_Direction, and Orientation as key determinants. The findings suggest that visibility, accessibility, and spatial configuration play a dominant role in occupant perception. Overall, the study highlights the potential of ML as a decision-support tool for evidence-based, occupant-centered healthcare design, particularly in dense urban contexts.

Funding: This research received no external funding

Acknowledgments: We would like to thank the concerned authorities for providing access to the school during data collection and for their necessary administrative support. We are also grateful to the field staff for their assistance in coordinating with the students.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Ulrich, R.S. View Through a Window May Influence Recovery from Surgery. Science (1979) 1984, 224, 420–421, doi:10.1126/SCIENCE.6143402.
2. Andrade, C.; Lima, M.L.; Fornara, F.; Bonaiuto, M. Users' Views of Hospital Environmental Quality: Validation of the Perceived Hospital Environment Quality Indicators (PHEQIs). J Environ Psychol 2012, 32, 97–111, doi:10.1016/J.JENVP.2011.12.001.
3. Frontczak, M.; Wargocki, P. Literature Survey on How Different Factors Influence Human Comfort in Indoor Environments. Build Environ 2011, 46, 922–937, doi:10.1016/J.BUILDENV.2010.10.021.
4. Zaman, S.U.; Yesmin, M.; Pavel, M.R.S.; Jeba, F.; Salam, A. Indoor Air Quality Indicators and Toxicity Potential at the Hospitals' Environment in Dhaka, Bangladesh. Environ Sci Pollut Res Int 2021, 28, 37727–37740, doi:10.1007/S11356-021-13162-8.
5. Pati, D.; Harvey, T.E.; Willis, D.A.; Pati, S. Identifying Elements of the Health Care Environment That Contribute to Wayfinding. HERD: Health Environments Research & Design Journal 2015, 8, 44–67, doi:10.1177/1937586714568864.
6. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining 2016, 13-17-August-2016, 785–794, doi:10.1145/2939672.2939785/SUPPL_FILE/KDD2016_CHEN_BOOSTING_SYSTEM_01-ACM.MP4.
7. Lundberg, S.M.; Lee, S.I. A Unified Approach to Interpreting Model Predictions. Adv Neural Inf Process Syst 2017, 2017-December, 4766–4775.