Predicting the Thermal Comfort of Rickshaw Pullers in Outdoor Settings Using Machine Learning Approach

Humayun Kabir, Md. Hasan*, Md Arif Hasan, Abdullah-Al-Mamun, Abdullah Al Masud

Building Energy & Environmental Management-BEEM Enhanced by AI, Dhaka, Bangladesh
*21306026@iubat.edu

Abstract. Ensuring thermal comfort is crucial for rickshaw pullers, particularly in densely populated urban areas like Dhaka, Bangladesh, where extreme environmental conditions can negatively impact their health and productivity [1]. These workers are exposed to high temperatures, humidity, and pollution levels for prolonged periods, making them highly vulnerable to thermal stress [2-3]. As outdoor workers, they face challenges in managing their comfort, which directly affects their performance and well-being [4-5]. This study employs machine learning (ML) techniques to predict the thermal comfort of rickshaw pullers in urban settings, providing valuable insights for enhancing outdoor working conditions. A total of 600 data samples were collected through field surveys and environmental monitoring, capturing variables such as temperature, humidity, CO2 levels, lighting intensity, and physiological factors like body mass index (BMI) and weight. The study utilized three ML classifiers Random Forest (RF), Decision Tree (DT), and XGBoost (XGB) and evaluated their performance using metrics such as accuracy, Mean Squared Error (MSE), and Root Mean Squared Error (RMSE). SHapley Additive exPlanations (SHAP) analysis was applied to interpret feature importance and improve model transparency. All models achieved an accuracy of over 70%, with SHAP analysis revealing that temperature, weather quality, BMI, and weight were the most influential factors affecting thermal comfort. These findings highlight the potential of interpretable ML models to offer actionable insights that can improve outdoor working conditions. Future research should expand the dataset to include seasonal variations and additional features, which could refine predictive models and inform adaptive design strategies for outdoor environments.

Keywords: Thermal Comfort, Machine Learning, Urban Area, Thermal Stress, Seasonal Variation.

1 Introduction

Thermal comfort in outdoor environments is a critical aspect of occupational well-being, particularly for urban informal workers such as rickshaw pullers, who are regularly exposed to extended periods of heat stress in their daily activities [6-7]. This directly affects their physical health, mental alertness, and overall work performance. For rickshaw pullers, who operate under intense solar radiation, fluctuating weather conditions,

R. Rzayev et al. (eds.), *Proceedings of the International Conference on Current Problems in Engineering and Applied Sciences (ICCPEAS 2025)*, Advances in Engineering Research 299,
https://doi.org/10.2991/978-94-6239-668-5_91

and heavy traffic exposure, discomfort can lead to fatigue, reduced productivity, and increased vulnerability to heat-related illnesses [8]. In densely populated cities like Dhaka, Bangladesh, these challenges are intensified due to high ambient temperatures, limited shaded routes, and inadequate urban infrastructure that fails to support thermal relief. As climate change continues to exacerbate extreme heat events, understanding and predicting the thermal comfort of rickshaw pullers becomes essential for developing targeted interventions, safeguarding health, and improving urban mobility systems. To be able to effectively quantify thermal stress and promote evidence-based planning for outdoor occupational environments, this emphasizes the necessity of data-driven, adaptable approaches like machine learning algorithms [9].

Research on thermal comfort has mostly concentrated on indoor settings, leaving outdoor occupational situations especially those involving low-income workers largely understudied, despite the growing concern over heat stress in metropolitan areas [10]. Rickshaw pullers are a particularly vulnerable population whose thermal comfort and health hazards are frequently disregarded in current studies because they work long hours on the road in intense sunshine, high humidity, and heavy traffic [11]. The intricate relationships between temperature, humidity, wind speed of pulling a rickshaw are not adequately captured by conventional assessment techniques like subjective surveys or single-parameter environmental measures. Furthermore, very little research has looked at how outdoor microclimatic conditions impact the productivity, well-being, and day-to-day working experience of informal laborer's in crowded urban areas like Dhaka. This reinforces the need for data-driven approaches, which include machine learning, to accurately predict thermal comfort and support evidence-based solutions for this impoverished group. It also draws attention to a significant knowledge gap [12].

The primary objective of this study is to develop and evaluate ML-based models for forecasting rickshaw pullers' thermal comfort in outdoor urban settings. Data was gathered using structured surveys and on-site measurements. To determine the most dependable prediction strategy, model performance was evaluated using metrics such as accuracy, MSE, and RMSE. The proportionate impact of each environmental and individual factor on thermal comfort levels was also assessed using SHapley Additive exPlanations (SHAP) analysis [13]. The findings aim to support the creation of data-driven interventions that can enhance the well-being of vulnerable outdoor workers. Additionally, this study provides a foundation for future research focused on improving urban heat-stress adaptation strategies.

2 Literature Review

Research on thermal comfort has traditionally centered on controlled indoor environments, where factors such as air temperature, humidity, and metabolic rate can be systematically assessed. Classical models such as Fanger's Predicted Mean Vote (PMV) and related thermal indices have guided much of this work [14]. However, scholars increasingly argue that these models fail to capture the complexity of outdoor environmental conditions, particularly in tropical and subtropical climates where temperature fluctuations, solar radiation, and wind speed vary throughout the day. Growing concerns about climate change and urban heat islands have led to greater interest in understanding outdoor thermal comfort, especially for vulnerable groups exposed to extreme temperatures. This research highlighted the growing but still limited body of research addressing how outdoor microclimates shape the thermal experiences of low-income workers in densely populated urban contexts [15].

Traditional thermal comfort models provide foundational knowledge; they are often inadequate for evaluating real-world outdoor environments. Studies conducted in hot climates consistently show discrepancies between model-predicted comfort levels and workers' actual thermal perceptions [16]. Research focusing on occupational heat exposure reveals significant physiological consequences for individuals who work outdoors for extended periods, such as farmers, construction laborers, and street vendors. Yet rickshaw pullers a group uniquely exposed to heat due to their prolonged physical exertion and continuous movement through congested city roads have received limited academic attention. Existing studies demonstrate that prolonged exposure to high temperatures and humidity contributes to fatigue, cardiovascular strain, dehydration, and reduced work performance among outdoor workers. However, the daily microclimatic variations encountered by rickshaw pullers remain largely unmeasured and underreported [17]. Moreover, research also highlights the limitations of single-parameter assessments, such as relying solely on temperature or humidity, which cannot capture the non-linear interactions that define heat stress in dynamic urban landscapes.

the growing application of ML in predicting thermal comfort across different environments. Researchers have demonstrated that ML models such as Random Forest, Decision Tree, Support Vector Machines, Artificial Neural Networks, and XGBoost provide more accurate predictions than traditional thermal comfort indices. These models excel at recognizing complex relationships among environmental variables, physiological attributes, and subjective comfort responses. Several studies also emphasize the importance of model interpretability, advocating for tools such as SHAP to evaluate the contribution of each feature in prediction tasks. Despite this advancement, most ML-driven thermal comfort studies focus on indoor scenarios like offices, classrooms, and semi-outdoor spaces such as courtyards. Outdoor occupational settings, particularly among informal workers, remain understudied [18]. The limited research that exists tends to concentrate on agricultural or industrial workers, overlooking transport-based laborers such as rickshaw pullers who experience continuous exposure to environmental stressors while performing intense physical activity.

The effects of heat on worker health and productivity, there is little empirical evidence capturing the environmental realities of rickshaw pullers in megacities like Dhaka. The combination of harsh microclimatic conditions intense solar radiation, high humidity, traffic emissions, poor air quality, and limited shade creates a unique thermal environment not adequately represented in existing studies [19]. This gap underscores the need for data-driven approaches that integrate environmental measurements with individual physiological characteristics to predict thermal comfort more accurately. ML models offer a promising avenue for advancing this understanding, especially when paired with interpretability tools that clarify the influence of key factors. By building on existing research while addressing overlooked populations, studies that apply ML to outdoor thermal comfort among rickshaw pullers can provide valuable insights for urban planning, worker protection, and climate adaptation strategies.

3 Methodology

This study followed a structured and comprehensive methodology as shown in figure-1 to predict the thermal comfort of rickshaw pullers operating in outdoor urban environments using ML algorithms, as illustrated in the overall research framework. The investigation was conducted within densely populated metropolitan zones of Dhaka, Bangladesh. Data collection was carried out through structured field surveys and direct sensor-based environmental monitoring during active working hours, specifically between 9:00 AM and 4:00 PM, to capture diurnal variations in temperature, humidity, and air quality. A total of 600 valid data samples were gathered, ensuring a balanced representation of varying locations, traffic conditions, and individual demographics. The dataset comprised a comprehensive set of features reflecting the multidimensional aspects of outdoor thermal comfort. The environmental parameters included temperature (°C), relative humidity (%), CO_2 concentration (ppm), noise level (dB), lighting level (lux), total volatile organic compounds (TVOC), wind speed (km/hr). Physiological and personal attributes encompassed weight (kg), age, Body Mass Index (BMI), clothing type, and indicators of drug addiction or substance use. Socioeconomic and occupational factors were also considered, such as profession duration (years), daily working hours, daily earning capacity, dwelling place, family member size, and education level. Structural and lifestyle elements like structural class of the vehicle and loan status were recorded to reflect the broader living and working context of the respondents. Each entry in the dataset represented a unique combination of environmental exposure and individual condition, associated with a subjective comfort rating on a categorical thermal comfort scale which served as the target variable for prediction. Raw data were initially inspected for missing or inconsistent values, which were addressed through imputation or removal depending on the variable type and degree of absence.

The environmental parameters included temperature (°C), relative humidity (%), CO_2 concentration (ppm), noise level (dB), lighting level (lux), total volatile organic compounds (TVOC), wind speed (km/hr). Physiological and personal attributes encompassed weight (kg), age, Body Mass Index (BMI), clothing type, and indicators of drug addiction or substance use. Socioeconomic and occupational factors were also considered, such as profession duration (years), daily working hours, daily earning capacity, dwelling place, family member size, and education level. Structural and lifestyle elements like structural class of the vehicle and loan status were recorded to reflect the broader living and working context of the respondents. Each entry in the dataset represented a unique combination of environmental exposure and individual condition, associated with a subjective comfort rating on a categorical thermal comfort scale which

served as the target variable for prediction. Raw data were initially inspected for missing or inconsistent values, which were addressed through imputation or removal depending on the variable type and degree of absence.

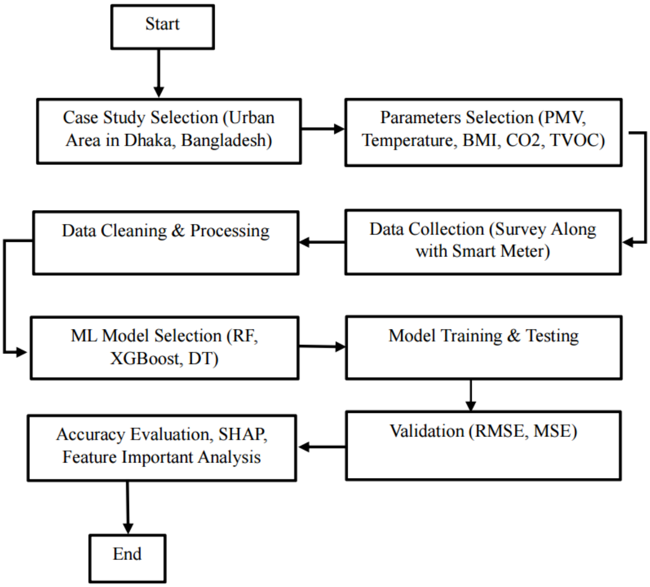


Fig1. Flow Chart of Research Methodology

Three ML classifiers DT, RF, and XGBoost were implemented to predict the comfort state of rickshaw pullers. These algorithms were chosen due to their high efficiency in handling nonlinear interactions, categorical variables, and mixed data structures. The dataset was partitioned into 80% training and 20% testing subsets to evaluate model generalization, and hyper parameters for each classifier were optimized using GridSearchCV with cross-validation to identify the best-performing configurations. The DT model was selected for its interpretability, the RF for its ensemble-based robustness and capability to reduce overfitting, and XGBoost for its efficiency in capturing complex and nonlinear relationships between variables.

SHAP analysis was employed after model training. SHAP provides a consistent approach to quantify the individual contribution of each feature to the final model output, thereby allowing a deeper understanding of the determinants influencing thermal comfort. Model performance was evaluated using classification accuracy, MSE, and RMSE, providing quantitative measures of predictive reliability and error distribution. The formulas for MSE and RMSE were as follows [20].

$$MSE = \frac{1}{m} \sum_{i=1}^{m} (X_i - Y_i)^2 \tag{1}$$

$$RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^{m} (X_i - Y_i)^2} \tag{2}$$

Here, Xi represents the predicted i-th value.
Yi represents the actual i-th value.
The regression method predicts the Xi value for the corresponding Yi value in the ground truth dataset. Two constants are defined as the mean of the true values. These metrics allowed the study to determine each model's classification performance as well as the size of prediction errors, guaranteeing that the models assigned were accurate and reliable across all conditions. This methodological framework integrates multi-dimensional data processing, interpretable ML modeling, and explainable feature analysis to predict the thermal comfort of outdoor laborers. By systematically incorporating environmental, physiological, and occupational variables, the approach ensures both accuracy and interpretability in comfort prediction.

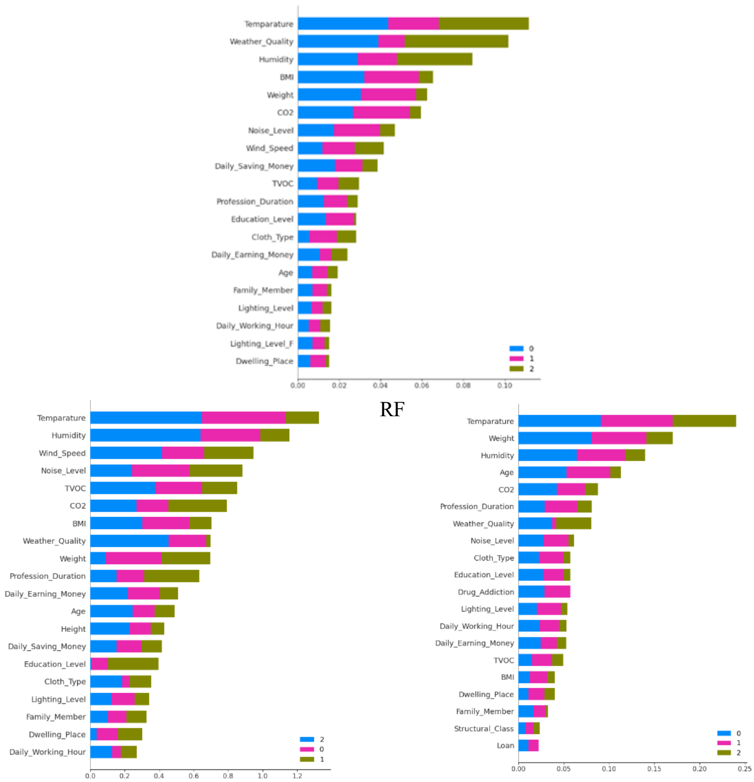
4. Results and Discussion

Based on the comparative analysis of three ML models such as DT, XGBoost, and RF for predicting student comfort level as influenced by sitting surfaces, notable performance variations were observed across the models. The data in Table-1 shows that, the RF model obtained the highest overall predictive performance, showing an accuracy of 0.75, a MSE of 0.57, and RMSE of 0.75. This indicates that the RF model accurately identified the fundamental relationships between environmental, demographic, and physical features influencing comfort perception in outdoor campus environments. XGBoost also showed competitive performance, recording an accuracy of 0.73, MSE of 0.68, and RMSE of 0.82, reflecting strong predictive capability but slightly reduced accuracy compared to RF. On the other hand, the DT model showed relatively lower results, with an accuracy of 0.71, MSE of 0.70, and RMSE of 0.84, showing limited effectiveness in modeling comfort levels from the given feature set. Although both DT and RF are tree based models, the accuracy variation between them was notably high.

Table 1. ML Model Performance

Parameters	**ML Model**	**Accuracy**	**MSE**	**RMSE**
PMV level	**RF**	**0.75**	**0.57**	**0.75**
	XGBoost	**0.73**	**0.68**	**0.82**
	DT	**0.71**	**0.70**	**0.84**

The SHAP analysis in Fig. 2. further identified the key variables influencing comfort perception, showing that Temperature is the most significant predictor in all models, highlighting its direct physiological and perceptual effects on thermal perception during extended outside work. Humidity, CO_2, noise level, wind speed, and weather quality similarly show up as influential across models, suggesting that comfort responses are simultaneously shaped by both larger climatic circumstances and micro environmental air-quality characteristics. Individual thermoregulatory variations are reliably reflected in model outputs by body-related parameters including age, weight, and BMI; however, their influence is notably greater in the Random Forest and XGBoost models than in the more straightforward Decision Tree. The tenure of a profession, daily earnings and savings, and daily working hours are examples of socioeconomic and work-pattern characteristics that show moderate influence across all models, indicating that long-term occupational exposure and lifestyle determinants interact slightly with physiological strain. Because of its more rigid structure, the Decision Tree produces more compressed significance values, whereas XGBoost typically assigns higher SHAP magnitudes across variables, demonstrating its greater sensitivity and ability to capture non-linear correlations. It's interesting to see that environmental contaminants like CO2 and TVOC score higher in Random Forest and XGBoost, emphasizing the importance of air quality in outdoor thermal stress a feature that the Decision Tree less effectively captures. Overall, the comparative findings show that ensemble models offer a more nuanced understanding of thermal comfort determinants, supporting the combined significance of occupational conditions, personal characteristics, and microclimate in predicting rickshaw pullers' comfort in actual outdoor settings.



XGBoost DT

Fig. 2. Shap Diagram of RF, XGBoost & DT Models

5. Conclusion

A comparison of the Random Forest, XGBoost, and Decision Tree models shows that machine learning methods provide a reliable framework for forecasting the thermal comfort of rickshaw pullers operating in difficult outdoor conditions. All three models consistently identify temperature as the most important factor influencing thermal comfort, despite variations in model complexity and sensitivity. This highlights temperature's paramount significance in influencing physiological and perceptual reactions during intense physical exercise. Thermal comfort is a multifaceted outcome influenced by both atmospheric conditions and micro-environmental stressors, as evidenced by the significant contributions of other environmental variables, particularly humidity, wind speed, CO2 concentration, noise level, and overall weather quality. The significance of individual adaptability and thermoregulatory traits is further highlighted by individual parameters including BMI, weight, age, and long-term professional exposure, indicating that comfort perceptions differ not only with environment but also with physical and vocational profiles. While RF provides consistent and comprehensible patterns of varying relevance, XGBoost exhibits the highest sensitivity among the models in recognizing intricate, non-linear interactions among predictors. Despite its simplicity, the DT offers a clear hierarchical picture of feature influence, but with less detail. When taken as a whole, the results demonstrate the usefulness of ensemble learning techniques for producing subtle insights that can assist focused treatments. In the end, these models not only advance scientific knowledge of outdoor thermal comfort but also offer practical evidence to guide policy, occupational health recommendations, and urban design strategies meant to protect outdoor workers' comfort and productivity in increasingly unpredictable weather.

Acknowledgement
No specific grant for a public, private, or nonprofit funding organization was obtained for this study.

Conflict of Interest
The authors declare that they have no conflict of interest.

Data Availability
Data is Available on request from the corresponding author.

References

1. Wadood, S.N., & Tehsum, M. (2018). Examining Vulnerabilities: the Cycle Rickshaw Pullers of Dhaka City.
2. Sharmin, T., & Steemers, K. (2018). Effects of microclimate and human parameters on outdoor thermal sensation in the high-density tropical context of Dhaka. International Journal of Biometeorology, 64(2), 187–203. https://doi.org/10.1007/s00484-018-1607-2

3. Burkart, K., Schneider, A., Breitner, S., Khan, M. H., Krämer, A., & Endlicher, W. (2011). The effect of atmospheric thermal conditions and urban thermal pollution on all-cause and cardiovascular mortality in Bangladesh. Environmental Pollution, 159(8–9), 2035–2043. https://doi.org/10.1016/j.envpol.2011.02.005
4. Abokhashabah, T., Jamoussi, B., Summan, A. S., Abdelfattah, E., & Ahmad, I. (2021). Effects of heat stress exposure and climate change on health and safety of outdoor workers. Current World Environment, 16(3), 836–847. https://doi.org/10.12944/cwe.16.3.15
5. Venugopal, V., Shanmugam, R., & Kamalakkannan, L. P. (2021). Heat-health vulnerabilities in the climate change context comparing risk profiles between indoor and outdoor workers in developing country settings. Environmental Research Letters, 16(8), 085008. https://doi.org/10.1088/1748-9326/ac1469
6. X. Yao, Y. Ping, Y. (Wolf) Song, and P. Vink, "Sitting comfort in an aircraft seat with different seat inclination angles," Int J Ind Ergon, vol. 96, p. 103470, Jul. 2023, doi: 10.1016/J.ERGON.2023.103470.
7. S. Liu, P. Beillas, L. Ding, and X. Wang, "PIPER adult comfort: an open source full body human body model for seating comfort assessment and its validation under static loading conditions," Front Bioeng Biotechnol, vol. 11, p. 1170768, May 2023, doi: 10.3389/FBIOE.2023.1170768/BIBTEX.
8. Begum, S., & Sen, B. (2004). Unsustainable livelihoods, health shocks and urban chronic poverty: Rickshaw Pullers as a case study. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.1754402
9. Jeong, J., Jeong, J., Lee, M., Lee, J., & Chang, S. (2022c). Data-driven approach to develop prediction model for outdoor thermal comfort using optimized tree-type algorithms. Building and Environment, 226, 109663. https://doi.org/10.1016/j.buildenv.2022.109663
10. Hashemi, F., Poerschke, U., Iulo, L. D., & Chi, G. (2023). Urban Microclimate, Outdoor Thermal Comfort, and Socio-Economic Mapping: A case study of Philadelphia, PA. Buildings, 13(4), 1040. https://doi.org/10.3390/buildings13041040
11. Lal, P., Batra, M., Majhi, M. M., Ahuja, P., & Bhatnagar, N. (2022b). Occupational health hazards of rickshaw pullers in lower middle income Country, India. Indian Journal of Occupational and Environmental Medicine, 26(4), 281–284. https://doi.org/10.4103/ijoem.ijoem_113_21
12. El-Khalik, W. A. (2022). A machine learning approach for improved thermal comfort prediction in sustainable built environments. Sustainable Machine Intelligence Journal, 1. https://doi.org/10.61185/smij.2022.11101
13. Lan, H., Hou, H., & Gou, Z. (2023). A machine learning led investigation to understand individual difference and the Human-Environment interactive effect on classroom thermal comfort. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.4334325
14. Niza, I. L., & Broday, E. E. (2022). Development of thermal comfort models over the past years: a systematic literature review. International Journal of Ambient Energy, 43(1), 8830–8846. https://doi.org/10.1080/01430750.2022.2116597

15. Banerjee, S., & Chattopadhyay, S. (2020). A meta-analytical review of outdoor thermal comfort research: Applications, gaps and a framework to assess low-income settlements in Indian megacities. Urban Climate, 33, 100641. https://doi.org/10.1016/j.uclim.2020.100641
16. Ramos, J. S., Toulou, A., Delgado, M. G., Amores, T. R. P., Medina, D. C., & Domínguez, S. Á. (2022). Thermal Resilience of Citizens: Comparison between Thermal Sensation and Objective Estimation in Outdoor Spaces: A Case Study in Seville, Spain. Applied Sciences, 12(22), 11676. https://doi.org/10.3390/app122211676
17. Runkle, J. D., Cui, C., Fuhrmann, C., Stevens, S., Del Pinal, J., & Sugg, M. M. (2019). Evaluation of wearable sensors for physiologic monitoring of individually experienced temperatures in outdoor workers in southeastern U.S. Environment International, 129, 229–238. https://doi.org/10.1016/j.envint.2019.05.026
18. Fujishiro, K. (2019). Informal employment as a frontier of occupational safety and health research. Industrial Health, 57(6), 653–654. https://doi.org/10.2486/indhealth.57_600
19. Johansson, E., & Emmanuel, R. (2006). The influence of urban design on outdoor thermal comfort in the hot, humid city of Colombo, Sri Lanka. International Journal of Biometeorology, 51(2), 119–133. https://doi.org/10.1007/s00484-006-0047-6
20. Chicco, D., Warrens, M. J., & Jurman, G. (2021). The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. PeerJ Computer Science, 7, e623. https://doi.org/10.7717/peerj-cs.623