

Machine Learning for Heatstroke Prediction: Unlocking Weather and Climate Signals

Jingyao Zhu^{1,2}

1) Pudong Foreign Language School is affiliated with SISU, Pudong New District, Shanghai, China

2) Cambridge Centre for International Research, Cambridge Science Park, Milton Road, Cambridge, UK; elenazhu2007@gmail.com

Mentor: Marco Vinicio Alban-Paccha

ABSTRACT: Heatstroke is a life-threatening condition that poses growing challenges under rising global temperatures. This study aimed to develop and evaluate machine learning models for predicting daily heatstroke incidence using nationwide health and meteorological data from Japan. Predictors included maximum, minimum, and average temperature, relative humidity, calculated heat index, calendar variables, and prefecture identifiers. Five modelling approaches were compared: Poisson regression, random forest, XGBoost, LightGBM, and LSTM. Model performance was assessed using mean squared error (MSE), root mean squared error (RMSE), and coefficient of determination (R^2). Among all models, LightGBM (MSE = 98.15, RMSE = 4.26, R^2 = 0.827) and XGBoost (MSE = 101.01, RMSE = 4.33, R^2 = 0.822) achieved better performance, outperforming both traditional Poisson regression (R^2 = 0.434) and neural network approaches (R^2 = 0.488). Feature importance analysis using SHapley Additive exPlanations (SHAP) confirmed maximum temperature as the strongest predictor. At the same time, the inclusion of prefecture identifiers substantially improved predictive power, underscoring the importance of regional heterogeneity. Excluding the prefecture reduced LightGBM performance to R^2 = 0.378. These findings highlight the potential of machine learning for early warning systems that integrate meteorological drivers and regional modifiers such as population. In the context of climate change and more frequent heatwaves, such adaptive models can support targeted public health strategies, optimize resource allocation, and reduce the health and economic burden of heatstroke.

KEYWORDS: Computational Biology and Bioinformatics, Computational Epidemiology, Heatstroke Prediction, Climate Change and Health, Machine Learning Models.

■ Introduction

Heatstroke is one of the most severe manifestations of heat-related illness, arising when the body's core temperature exceeds the compensatory capacity of thermoregulatory mechanisms. It can be triggered by environmental exposure to high temperatures, excessive physical exertion, or a combination of both, and typically presents as a spectrum of conditions ranging from mild heat exhaustion to potentially fatal heatstroke. The defining characteristics include a core body temperature exceeding 40 °C, central nervous system dysfunction (such as confusion, seizures, or coma), and progressive multi-organ failure. Without timely medical intervention, the case-fatality rate can exceed 50%, underscoring its urgency as a global public health issue.^{1,2}

The burden of heatstroke is not limited to individual health outcomes. It has significant societal implications, particularly for vulnerable populations and high-risk occupational groups such as outdoor workers, athletes, emergency responders, and the elderly (Figure 1). Epidemiological studies consistently demonstrate a strong association between elevated ambient temperatures and an increase in hospital admissions, ambulance calls, and mortality attributable to heatstroke.³ The Fire and Disaster Management Agency of Japan (FDMA) has reported a persistent upward trend in heatstroke-related emergency transports across the country. While annual fluctuations

occur due to the presence or absence of extreme heat events and variations in rainy season length, the long-term trajectory reflects a clear increase in heat-related morbidity. This pattern is projected to worsen in the coming decades due to climate change, with intensifying heatwaves posing escalating risks to public health, occupational safety, and healthcare systems worldwide.

Beyond morbidity and mortality, heatstroke also carries a substantial economic burden. Hospitalizations for severe cases require intensive care, mechanical ventilation, and prolonged rehabilitation, placing a significant strain on healthcare systems during the peak summer months. Similar patterns have been observed internationally, where rising temperatures significantly increase hospital admissions and healthcare costs attributable to heat exposure.^{4,5} The cost extends to productivity losses, as affected individuals may need extended recovery time away from work or experience long-term neurological and cardiovascular complications that reduce functional capacity. For employers, especially those in sectors reliant on outdoor labor, such as agriculture, construction, and transportation, heat-related illnesses contribute to absenteeism, reduced work performance, and occupational hazards. At a societal level, these direct and indirect costs combine to form a considerable economic impact, further justifying the urgent need for preventive strategies and predictive models.⁶⁻⁸

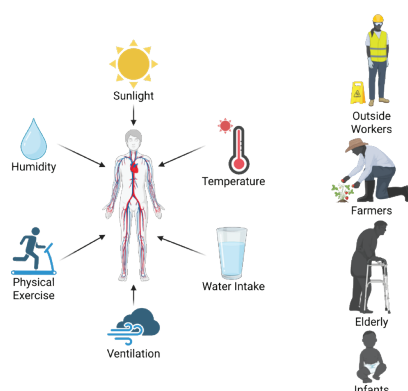


Figure 1: Environmental and behavioral factors contributing to heat stress (sunlight, temperature, humidity, sports, ventilation, and water intake) and the vulnerable groups most affected, including outdoor workers, farmers, the elderly, and infants.

Several studies illustrate this paradigm shift. Ogata *et al.* developed a machine learning model that combined multivariate meteorological and demographic data from 16 Japanese cities, achieving superior predictive accuracy over traditional Wet-Bulb Globe Temperature (WBGT)-based approaches for 12-hour incidence forecasting.⁹ Xu *et al.* conducted a comprehensive comparison of machine learning algorithms, including random forest and gradient boosting methods, in a southern Chinese city and demonstrated that ensemble models significantly outperformed generalized linear models, particularly in capturing nonlinearities.¹⁰ Shimazaki *et al.* advanced the field further by incorporating wearable sensor data (including predicted heat strain tolerance, heart rate, and activity level) into a supervised machine learning model that achieved 85.2 % predictive accuracy in occupational settings.¹¹ Meanwhile, Takada *et al.* demonstrated that simulated core body temperature could serve as a robust predictor of ambulance calls in Japanese metropolitan areas, performing comparably to environmental temperature-based models but with fewer input parameters.¹² Collectively, these studies highlight the growing utility of machine learning for predictive modelling of heat-related illness.

However, current literature has two key limitations. First, most studies have been restricted to single cities or a limited set of urban areas, thereby limiting the generalizability of their findings. Second, while meteorological and physiological features have been studied extensively, the potential contribution of geographical heterogeneity (e.g., regional socio-economic, demographic, and environmental differences) remains underexplored. This gap is particularly salient in a country like Japan, where diverse climatic zones, population structures, and levels of urbanization create region-specific vulnerabilities.

The present study addresses these gaps by conducting a nationwide analysis of heatstroke incidence across Japan using five years of health and meteorological data (2015–2019, June to September). By systematically comparing multiple machine learning approaches, including Poisson regression, random forest, XGBoost, LightGBM, and LSTM, the study evaluates predictive performance, identifies key features influencing outcomes, and assesses the role of regional variation in shaping heatstroke risk. In particular, the study applies SHapley Additive exPlanations (SHAP) to quantify feature importance,

enabling interpretable insights into the drivers of heatstroke incidence.

Through this work, we aim to provide robust evidence for the integration of machine learning models into early warning systems for heat-related illness. By highlighting both the meteorological and regional determinants of heatstroke, this study contributes to the development of tailored public health strategies that account for geographical heterogeneity, ultimately supporting more effective risk prediction, prevention, and resource allocation in the face of escalating climate change.

■ Methods

Iterative machine learning workflow:

This study employs an iterative machine learning workflow aimed at systematically enhancing model performance through continuous refinement (Figure 2). The process commences with data preparation, which encompasses data integration and preprocessing. Following this, informative features are constructed via feature engineering, and models are trained using the training set. The trained models are subsequently evaluated on a test set. Should the performance fail to meet predefined criteria, feedback loops facilitate a return to earlier stages for further refinement. This "training-evaluation-feedback" cycle continues iterating until the model fulfills the established performance standards. Finally, the trained model is evaluated on a held-out test set to assess predictive performance and generalizability.

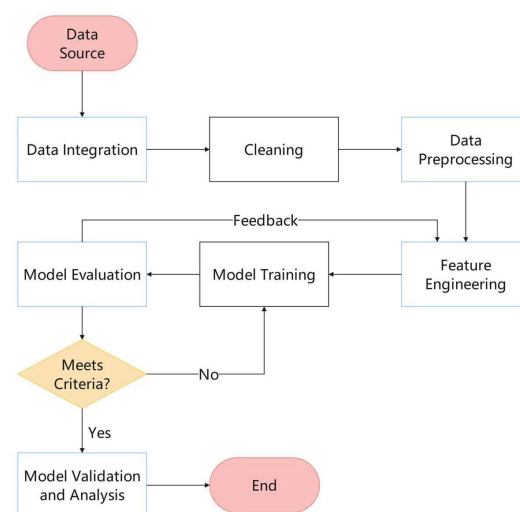


Figure 2: An iterative machine learning workflow guides the model development process, beginning with data sourcing and integration, moving systematically to cleaning, preprocessing, and feature engineering, and concluding with the model training phase. The model is then evaluated against predefined criteria, with feedback loops enabling cyclical refinement of data, features, or model parameters until performance standards are met.

Data sources:

This study utilized a publicly available dataset from Zenodo: *Datasets for quantifying association between short-term exposure to maximum temperature and heatstroke-related ambulance dispatches in Japan: a time-stratified case-crossover design* (<https://zenodo.org/records/14560225>), supplemented with data obtained from the official website of the Japan Meteorological Agency (JMA). The public dataset included daily records of

heatstroke-induced ambulance dispatches (HSAD) and meteorological data across Japan's 47 prefectures from 2015 to 2019, containing data of the summer months, June to September, when heatstroke incidence peaks. To enhance data completeness, daily minimum temperature and average temperature records were retrieved from the JMA website and integrated into the dataset. All data were sourced from official Japanese government channels, ensuring high reliability. The final dataset comprises 11 features and 28,200 records.

Study variables:

The outcome variable was the daily number of heatstroke cases per prefecture. Predictor variables included meteorological and temporal features:

- Daily maximum temperature (Tempmax, °C)
- Daily minimum temperature (Tempmin, °C)
- Daily average temperature (Tempave, °C)
- Average relative humidity (Rhumave, %)
- Calculated heat index (HeatIndex, composite variable of temperature and humidity)
 - Calendar information: year, month, day of month, day of week (Dow), and holiday status (binary indicator)
 - Prefecture identifier (categorical variable capturing regional variation)

To improve the predictive performance of the model, the heat index was added, a commonly used composite meteorological indicator. The Heat Index, also referred to as apparent temperature, is a meteorological measure that combines ambient temperature and relative humidity to reflect perceived thermal discomfort. Its primary function is to quantify the level of heat stress experienced by individuals under hot and humid conditions. High humidity impedes sweat evaporation, reduces the body's cooling efficiency, and thereby elevates the perceived temperature beyond the actual air temperature. As a result, the Heat Index is widely employed in public health, weather advisories, and occupational safety to assess heat-related illness risks, develop preventive strategies, and issue

$$HI = c_1 + c_2 \cdot T + c_3 \cdot RH + c_4 \cdot T \cdot RH + c_5 \cdot T^2 + c_6 \cdot RH^2 + c_7 \cdot T^2 \cdot RH + c_8 \cdot T \cdot RH^2 + c_9 \cdot T^2 \cdot RH^2$$

The Heat Index (HI) was calculated using the standard nine-coefficient Rothfus regression equation adopted by the U.S. National Weather Service. In this equation, HI denotes the Heat Index, T represents air temperature expressed in degrees Fahrenheit (°F), RH represents relative humidity (%), and the regression coefficients are defined as:

$$\begin{aligned} C_1 &= -42.379, \\ C_2 &= 2.04901523, \\ C_3 &= 10.14333127, \\ C_4 &= -0.22475541, \\ C_5 &= -6.83783 \times 10^{-3}, \\ C_6 &= -5.481717 \times 10^{-2}, \\ C_7 &= 1.22874 \times 10^{-3}, \\ C_8 &= 8.5282 \times 10^{-4}, \\ C_9 &= -1.99 \times 10^{-6}. \end{aligned}$$

Following the original National Weather Service formulation, temperature values were converted from degrees Celsius to degrees Fahrenheit before applying the regression equation.

Feature selection and explainability

To assess the relative importance of meteorological and temporal features, SHapley Additive exPlanations (SHAP) were applied. SHAP is a game-theoretic framework that attributes the contribution of each feature to the prediction outcome. A random forest regression model was used as the baseline for SHAP computation, chosen for its robustness in modelling nonlinear relationships and for the efficiency of the TreeExplainer algorithm within the SHAP library. The random forest was configured with 50 trees and a maximum depth of 5 to balance computational cost and prevent overfitting.

Data splitting strategy:

For the primary model comparison, the dataset was randomly divided into training and testing subsets. All models were trained using the same predictor variables and evaluated on identical train-test splits to ensure fair comparison across algorithms. To further assess robustness and address potential concerns regarding temporal dependence, an additional temporal hold-out validation was conducted. In this analysis, data from 2015–2018 were used for model training, while data from 2019 were reserved exclusively for testing.

Model development

Five modelling approaches were evaluated:

1. Poisson regression, a count model traditionally used in epidemiology for rare events.
2. Random forest regression is an ensemble of decision trees trained using bootstrap sampling and feature bagging.
3. Extreme Gradient Boosting (XGBoost) is a gradient boosting framework optimized for efficiency and performance.
4. Light Gradient Boosting Machine (LightGBM) is a gradient boosting framework using histogram-based algorithms for faster training.
5. Long Short-Term Memory (LSTM) neural network, a recurrent neural network architecture designed to capture sequential dependencies in time-series data.

All models were trained on the same input features and evaluated using identical training-test splits to ensure comparability. Hyperparameters for ensemble models were optimized using grid search. For LightGBM, the final configuration included a maximum of 31 leaves per tree, a learning rate of 0.05, 200 boosting iterations, and a random seed set to 42.

Model evaluation:

Model performance was evaluated using three standard metrics: Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and the Coefficient of Determination (R^2):

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$RMSE = \sqrt{MSE}$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

where n denotes the number of observations, y_i represents the observed heatstroke incidence for observation i, $\hat{y}_i$ represents

the corresponding model prediction, and $\bar{y}$ denotes the mean observed heatstroke incidence across all observations.

These metrics collectively capture the accuracy (MSE), interpretability (RMSE, in the same units as the outcome variable), and explanatory power (R^2) of the models. Comparisons across models were made to determine the most effective algorithm for predicting heatstroke incidence.

Ethical considerations:

All data used in this study were aggregated, anonymized, and publicly available. As no individual-level or identifiable health data were used, ethical approval was not required.

Results and Discussion

Descriptive analysis and observable patterns:

Between 2015 and 2019, a total of five years of nationwide data were analyzed for the summer months of June through September. Across all prefectures, heatstroke incidence showed marked temporal and spatial variation, with clear peaks during extreme heat events. Descriptive statistics indicated that meteorological variables, particularly temperature and humidity, were strongly associated with heatstroke incidence, whereas calendar-related features such as holiday status, year, and day of week had comparatively limited influence.

The calculated heat index demonstrated the strongest association with heatstroke counts among the meteorological variables, highlighting its utility as a composite indicator of thermal stress. Correlation analysis revealed coefficients of 0.48 for daily maximum temperature, 0.41 for minimum temperature, 0.47 for average temperature, and 0.51 for heat index, all with p -values ≤ 0.001 (Figure 3). These findings confirm that rising ambient temperatures and higher heat index values substantially increase the probability of heatstroke events. In contrast, variables such as year, month, day, and holiday status exhibited correlation coefficients below 0.1, suggesting their contribution was minor compared to direct meteorological drivers.

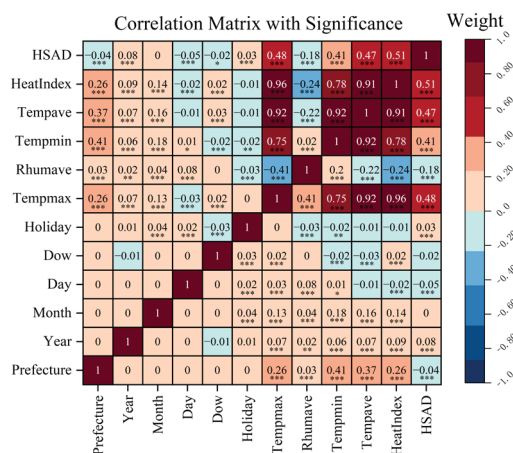


Figure 3: Correlation matrix with significance levels for meteorological, temporal, and regional features in relation to heatstroke admissions (HSAD). Strong positive correlations are observed between temperature variables and heat index, while weaker or negligible associations are seen for calendar-related features (year, month, day, holiday, and day of week).

Feature importance analysis:

To quantify the contribution of individual features, SHapley Additive exPlanations (SHAP) values were computed using a random forest model as the reference. The ranking of feature importance was: daily maximum temperature (Tempmax), prefecture identifier, daily average temperature (Tempave), heat index, daily minimum temperature (Tempmin), month, day, average relative humidity (Rhumave), year, and day of week (Dow). These results were used to rationalize and refine our feature set. Based on the SHAP-derived importance ranking, the study selected the six most influential predictors—daily maximum temperature (Tempmax), prefecture identifier, daily average temperature (Tempave), heat index, daily minimum temperature (Tempmin), and month—for the final model, ensuring it focuses on the dominant risk factors (Figure 4).

These results underscore two key findings. First, maximum temperature remains the single most powerful meteorological determinant of heatstroke incidence. Second, the inclusion of the prefecture identifier significantly improves model performance, pointing to strong regional heterogeneity. The prefecture variable effectively encodes differences in population structure, socio-economic conditions, urbanization level, and healthcare reporting standards, all of which modify the baseline vulnerability to heat-related illness.

Although relative humidity exhibited relatively low SHAP importance, this does not imply that humidity is unimportant for heatstroke risk. A likely explanation is that humidity information is already incorporated into the Heat Index variable. Because the heat index is calculated as a function of both temperature and relative humidity, substantial collinearity exists between these predictors. Consequently, part of the predictive contribution of humidity is captured by the Heat Index, reducing its standalone importance in the SHAP ranking.

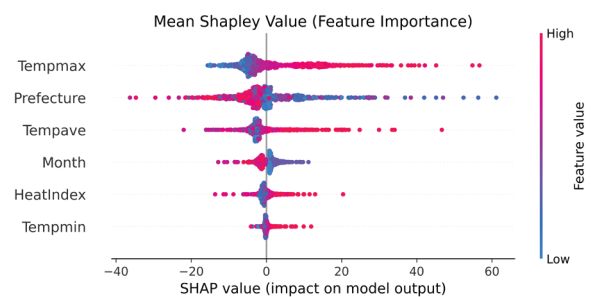


Figure 4: SHAP (SHapley Additive exPlanations) summary plot showing the relative importance of features in predicting daily heatstroke incidence. Maximum temperature (Tempmax) and prefecture identifier are the most influential predictors, followed by average temperature, heat index, and minimum temperature, while calendar-related features contribute minimally. The significant impact of the prefecture identifier suggests that localized factors, such as urban density or regional adaptation policies, critically modulate heatstroke risk beyond mere temperature. Consequently, this finding not only validates the model's design but also strongly advocates for the development of region-specific public health interventions.

Model comparison:

Five modelling approaches were evaluated: Poisson regression, random forest, XGBoost, LightGBM, and LSTM. Their performance was assessed using MSE, RMSE, and R^2 (Table 1).

Table 1: Predictive performance of machine learning and statistical models for daily heatstroke incidence, evaluated using mean squared error (MSE), root mean squared error (RMSE), and coefficient of determination (R^2). Notably, tree-based models, including XGBoost, Random Forest, and LightGBM, all achieved high explanatory power with R^2 values above 0.8, with LightGBM and XGBoost being the top performers. In contrast, both Poisson regression and LSTM demonstrated considerably lower predictive accuracy.

Model name	MSE	RMSE	R^2
Poisson regression	320.7344	7.4836	0.4340
XGBoost	101.0053	4.3259	0.8218
Random forest	108.6334	4.4663	0.8083
LSTM	290.4317	7.3702	0.4875
LightGBM	98.1498	4.2627	0.8268

Tree-based ensemble methods substantially outperformed the specific LSTM configuration evaluated in this study, as well as the traditional statistical model (Poisson regression). LightGBM and XGBoost achieved better overall performance. LightGBM achieved an MSE of 98.15 and an R^2 of 0.827, while XGBoost achieved an MSE of 101.01 and an R^2 of 0.822, representing a more than twofold improvement over Poisson regression. These results highlight the ability of gradient boosting algorithms to capture complex nonlinear interactions between meteorological factors and health outcomes. Because the performance differences among the top-performing models were relatively small, additional stability analyses were conducted using repeated random train-test splits to evaluate the robustness of these results.

Stability analysis across repeated random splits:

To evaluate the robustness of model performance across different random train-test partitions, the three best-performing machine-learning models were evaluated using 10 repeated random splits. The mean and standard deviation of R^2 values were calculated. LightGBM achieved a mean R^2 of 0.831 ± 0.017 , XGBoost achieved 0.828 ± 0.020 , and Random Forest achieved 0.809 ± 0.017 . The difference between LightGBM and XGBoost was substantially smaller than the observed run-to-run variability, indicating that the two algorithms exhibited comparable predictive performance. Both models consistently outperformed Random Forest.

Table 2: Mean and standard deviation (SD) of R^2 values calculated from 10 repeated random train-test splits using the same sampling strategy as the primary analysis.

Model name	Mean R^2	SD
LightGBM	0.831	0.017
XGBoost	0.828	0.020
Random forest	0.809	0.017

Temporal hold-out validation:

To evaluate the robustness of the findings under a more stringent forecasting scenario, an additional temporal hold-out validation was performed using data from 2015–2018 for training and data from 2019 for testing.

The results demonstrated that the overall ranking of models remained largely unchanged. LightGBM and XGBoost achieved the highest predictive performance ($R^2 = 0.769$), followed closely by Random Forest ($R^2 = 0.761$). Poisson Regression and LSTM yielded substantially lower predictive accuracy, with R^2 values of 0.433 and 0.415, respectively. Although these values were modestly lower than those obtained

using the random train-test split, the relative performance of the models remained consistent, indicating that the stronger performance of the tree-based ensemble models relative to the LSTM configuration evaluated in this study was robust to the choice of validation strategy.

Effect of regional heterogeneity:

The decisive role of regional characteristics was further examined through comparative experiments. When the prefecture feature was excluded, LightGBM performance dropped significantly ($R^2 = 0.378$). With the prefecture feature included, performance improved dramatically to $R^2 = 0.827$. This pattern was consistent across all models tested, demonstrating that regional differences account for a substantial proportion of the variance in heatstroke incidence (Figure 5).

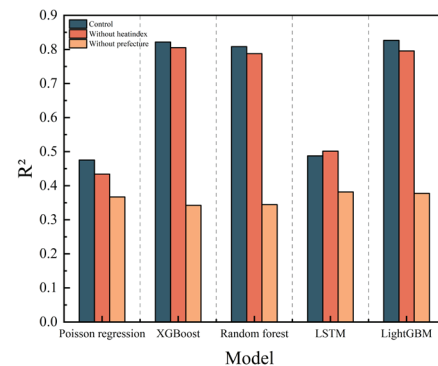


Figure 5: Comparison of predictive performance (R^2 values) across five models (Poisson regression, XGBoost, Random Forest, LSTM, and LightGBM) under three conditions: full model (control), without heat index, and without prefecture. The inclusion of prefecture notably improves accuracy across all models. The results reveal that the exclusion of the prefecture identifier significantly degraded model accuracy, whereas removing the heat index led to only a marginal decline.

Population adjustment analysis:

To examine whether the predictive contribution of the prefecture identifier primarily reflected differences in population size, an additional analysis was conducted by incorporating prefecture population as an explicit predictor.

The model, including meteorological variables and population, but excluding the prefecture identifier, achieved an R^2 of 0.850. This performance was comparable to that of the model including both population and prefecture information ($R^2 = 0.849$), suggesting that population size accounted for much of the predictive value associated with prefecture-level differences. These findings indicate that a substantial proportion of the regional variation observed in heatstroke case counts is related to population distribution across prefectures.

Interpretation of results:

LightGBM and XGBoost provided the most accurate predictions among the evaluated models and demonstrated strong robustness under both random train-test splitting and temporal hold-out validation. The models confirmed that meteorological drivers (maximum, minimum, and average temperature, as well as heat index) were the primary predictors, while regional differences, largely reflected by population distribution, provided additional predictive information.

This dual influence of weather and region implies that national-level warning systems may be insufficient if they do not account for localized vulnerabilities. For example, prefectures with higher elderly populations or stronger urban heat island effects may experience disproportionately greater impacts under identical meteorological conditions.

Implications:

The findings of this study highlight several important implications for the prediction and prevention of heatstroke. The strong predictive performance achieved by the machine-learning models suggests that routinely available meteorological information can be effectively incorporated into operational early-warning systems and public-health decision support. Such approaches may help identify periods of elevated risk and facilitate timely interventions to reduce heat-related health burdens.

Equally significant is the role of regional heterogeneity. The inclusion of the prefecture feature markedly improved predictive performance across all models, suggesting that heatstroke risk cannot be fully explained by meteorology alone. While prefecture-level information was initially identified as one of the most influential predictors, additional analyses suggest that a substantial portion of this effect is attributable to population size. Larger prefectures such as Tokyo and Osaka naturally record more heatstroke cases because more individuals are exposed to heat-related risks. Therefore, caution is warranted when interpreting the prefecture identifier as a direct proxy for vulnerability, urban heat-island intensity, or socio-economic characteristics. Future studies incorporating demographic, socioeconomic, and environmental indicators may help disentangle these mechanisms.

These results gain additional importance in the context of climate change. Rising global temperatures and the increasing frequency, intensity, and duration of heatwaves are expected to exacerbate the health burden of heatstroke worldwide. Projections for Japan and other temperate countries suggest that populations with little historical exposure to extreme heat will face growing risks in the coming decades.^{13,14} This challenge will be compounded by global economic impacts, where labor productivity losses and healthcare demand are projected to escalate across multiple sectors.^{15,16} Under these circumstances, adaptive and flexible prediction models will be essential. Machine learning frameworks, capable of integrating new environmental parameters and adjusting to shifting climate baselines, provide a pathway to resilient early warning systems.

Together, the evidence indicates that accurate heatstroke prediction requires the integration of both meteorological drivers and regional modifiers, embedded within a forward-looking framework that accounts for climate change. By incorporating these elements, machine learning models can underpin dynamic early warning systems that guide healthcare resource allocation, inform government policy, and deliver targeted interventions to protect vulnerable populations in an era of increasing thermal stress.

Future research directions:

Although substantial regional heterogeneity in heatstroke risk was observed, the drivers of these spatial disparities require further investigation. Future studies should integrate demographic, socioeconomic, urbanization, and climate projection data to develop region-specific prediction models and assess future heatstroke burden. External validation is also needed to facilitate implementation in local heat-health early warning systems. Furthermore, while this study evaluated established machine learning algorithms, future research could explore modeling frameworks that explicitly account for spatial heterogeneity and regional interactions to further improve predictive performance and interpretability.

■ Conclusion

This study systematically analyzed the relationship between meteorological conditions and heatstroke incidence across Japan, using five years of nationwide data. By comparing multiple modelling approaches, we demonstrated that LightGBM and XGBoost achieved the strongest predictive performance for nationwide heatstroke forecasting using meteorological and regional information. The findings confirm that machine-learning approaches can effectively capture the complex relationships underlying heatstroke occurrence and provide a reliable basis for predictive modelling. These findings emphasize that both environmental exposure and local socio-geographical factors must be considered to achieve accurate and actionable predictions.

The results carry important implications for public health. By capturing complex nonlinear interactions, machine learning models provide a reliable foundation for early warning systems capable of anticipating spikes in heatstroke cases. More importantly, the strong effect of regional variation underscores the necessity of tailoring preventive strategies to local conditions rather than relying solely on national averages. Such approaches can enable healthcare providers and government agencies to allocate resources more efficiently and deliver targeted guidance to high-risk populations.

Looking forward, the importance of accurate prediction systems will only grow under the pressures of climate change. Rising global temperatures and more frequent heatwaves will amplify the health burden of heat-related illness, particularly in populations with limited adaptive capacity. Prediction models must therefore not only integrate meteorological and socio-demographic data but also remain adaptable to shifting climatic baselines. The development of dynamic, multi-scale systems that combine environmental monitoring, demographic insights, and machine learning prediction represents a critical step toward building climate-resilient public health infrastructure.

In conclusion, this study provides evidence that machine learning, when combined with meteorological and regional features, can serve as a powerful tool for forecasting heatstroke risk. By embedding these insights into flexible, climate-aware early warning platforms, it is possible to mitigate the health impacts of extreme heat, protect vulnerable populations, and

reduce the societal costs of climate-driven thermal stress, which are expected to rise sharply under future warming scenarios.¹⁴⁻¹⁷

■ Acknowledgments

The author wishes to sincerely thank Dr Marco Alban-Paccha for his mentorship, encouragement, and invaluable guidance throughout this research journey. His support not only strengthened the project but also deepened the author's passion for engineering solutions to real-world challenges.

■ References

1. Armed Forces Health Surveillance Division. Update: Heat illness, active component, U.S. Armed Forces, 2021. *MSMR* 2022, 29 (5), 2–8. <https://health.mil/Reference-Center/Publications/2022/05/01/MSMR-Vol-29-No-5>.
2. Song, Q.; Mao, H. D.; Liu, S. Y. Definition and classification of heat illness. *Med. J. Chin. People's Liberation Army* 2019, 44 (7), 541–545. <https://doi.org/10.11855/j.issn.0577-7402.2019.07.01>.
3. Professional Committee of Critical Care Medicine of the Chinese PLA. Expert consensus on standardized diagnosis and treatment for heat stroke (draft). *Med. J. Chin. People's Liberation Army* 2015, 40 (1), 1–7. <https://doi.org/10.11855/j.issn.0577-7402.2015.01.01>.
4. Liao, S., Pan, W., Wen, L., Chen, R., Pan, D., Wang, R., Hu, C., Duan, H., Weng, H., Tian, C., Kong, W. Temperature-related hospitalization burden under climate change. *Nature*. 2025, 644, 960–968. <https://doi.org/10.1038/s41586-025-09352-w>.
5. Tong, M.X., Wondmagegn, B.Y., Williams, S., Hansen, A., Dear, K., Pisaniello, D., Xiang, J., Xiao, J., Jian, L., Scalley, B., Nitschke, M. Hospital healthcare costs attributable to heat and future projections. *Adv. Clim. Change Res.* 2021, 12 (2), 174–182. <https://doi.org/10.1016/j.accre.2021.07.008>.
6. Sapari, H., Selamat, M.I., Isa, M.R., Ismail, R., and Mahiyuddin, W.R.W. The Impact of Heat Waves on Health Care Services in Low- or Middle-Income Countries: Protocol for a Systematic Review. *JMIR Res Protoc.* 2023, 20. <https://pmc.ncbi.nlm.nih.gov/articles/PMC10616749>.
7. Chen, F., Zhang, X., and Chen, Z. Behind climate change: Extreme heat and health cost. *Structural Change and Economic Dynamics.* 2023, 64, 101–110. <https://doi.org/10.1016/j.strueco.2022.12.007>.
8. Boudreault, J.; Campagna, C.; Lavigne, E.; Chebana, F. Projecting the overall heat-related health burden and associated economic costs in a climate change context in Quebec, Canada. *Sci. Total Environ.* 2025, 958, 178022. <https://doi.org/10.1016/j.scitotenv.2024.178022>.
9. Ogata, S.; Takegami, M.; Ozaki, T.; Nakashima, T.; Onozuka, D.; Murata, S.; Nakaoku, Y.; Suzuki, K.; Hagihara, A.; Noguchi, T.; Iihara, K.; Kitazume, K.; Morioka, T.; Yamazaki, S.; Yoshida, T.; Yamagata, Y.; Nishimura, K. Heatstroke predictions by machine learning, weather information, and an all- population registry for 12-hour heatstroke alerts. *Nat. Commun.* 2021, 12, 4575. <https://doi.org/10.1038/s41467-021-24823-0>.
10. Xu, H.; Guo, S.; Shi, X.; Wu, Y.; Pan, J.; Gao, H.; Tang, Y.; Han, A. Machine learning-based analysis and prediction of meteorological factors and urban heatstroke diseases. *Front. Public Health* 2024, 12, 1420608. <https://doi.org/10.3389/fpubh.2024.1420608>.
11. Shimazaki, T.; Anzai, D.; Watanabe, K.; Nakajima, A.; Fukuda, M.; Ata, S. Heat stroke prevention in hot specific occupational environment enhanced by supervised machine learning with personalized vital signs. *Sensors* 2022, 22 (1), 395.
12. Takada, A.; Kodera, S.; Suzuki, K.; Nemoto, M.; Egawa, R.; Takizawa, H.; Hirata, A. Estimation of the number of heat illness patients in eight metropolitan prefectures of Japan: Correlation with ambient temperature and computed thermophysical responses. *Front. Public Health* 2023, 11, 1061135. <https://doi.org/10.3389/fpubh.2023.1061135>.
13. Anderson, G. B.; Bell, M. L.; Peng, R. D. Methods to calculate the heat index as an exposure metric in environmental health research. *Environ Health Perspect.* 2013, 121:1111–9. <https://doi.org/10.1289/ehp.1206273>.
14. Oyama, T.; Takakura, J.; Ishizaki, N. N.; Oka, K.; Honda, Y.; Masago, Y.; Hijioka, Y. Nationwide high- resolution heat risk projections and intervention cost analysis for the elderly in Japan under climate and demographic changes. SSRN preprint 2025. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5386992.
15. Ueta, H., Kodera, S., Sugimoto, S., and Hirata, A. Projection of future heat-related morbidity in three prefectures of Japan under 2 °C global warming scenarios. *Environ. Res.* 2024, 247, 118202. <https://doi.org/10.1016/j.envres.2024.118202>.
16. Sun, Y., Zhu, S., Wang, D., Duan, J., Lu, H., Yin, H., Tan, C., Zhang, L., Zhao, M., Cai, W., and Wang, Y. Global supply chains amplify economic costs of future extreme heat risk. *Nature* 2024, 627, 123– 129. <https://doi.org/10.1038/s41586-024-07147-z>.
17. Falchetta, G., De Cian, E., Sue Wing, I., and Carr, D. Global projections of heat exposure of older adults. *Nat. Commun.* 2024, 15, 47197. <https://doi.org/10.1038/s41467-024-47197-5>.

■ Author

Jingyao Zhu is a 12th-grade student from Pudong Foreign Language School affiliated with SISU, in Shanghai, China, who is passionate about engineering for social good. She developed LifeTag, a smart garment with sensors to monitor sanitation workers' health under extreme heat. Combining physics knowledge, self-taught programming, and innovation contest experience, she plans to pursue engineering at university, applying data science and machine learning to global health challenges.