

Explainable Artificial Intelligence Models for Predicting Malaria Risk in Kenya


Dennis Kariuki Muriithi ^{1,*}, Victor Wandera Lumumba , Olushina Olawale Awe ,
and Daniel Mwangi Muriithi 
Emails: dkariuki@chuka.ac.ke

ABSTRACT

The article aims to develop interpretable Machine Learning models using R statistical programming language for malaria risk prediction in Kenya, emphasizing leveraging Explainable AI (XAI) techniques to support targeted interventions and improve early detection mechanisms. The methodology involved using synthetic data with 1000 observations, employing over-sampling to address class imbalance, utilizing two machine learning algorithms (Random Forest and Extreme Gradient Boosting), applying cross-validation techniques, Hyper-parameter tuning and implementing feature importance and SHAP (Shapley Additive Explanations) for model interpretability. The findings revealed that Random Forest outperformed Extreme Gradient Boosting with 98% accuracy. Critical prediction features included clinical symptoms such as nausea, muscle aches, and fever, plasmodium species identification, and environmental factors like rainfall and temperature. Both models demonstrated strong sensitivity in detecting malaria cases. This promotes trust in model predictions by clearly outlining the decision process for individual outcomes. The research concluded that integrating Explainable AI into malaria risk prediction represents a transformative approach to public health management. Through providing transparent, interpretable models, the research offers a robust, data-driven approach to predicting malaria risks, potentially empowering healthcare providers and policymakers to deploy resources more effectively and reduce the disease burden in endemic regions.

Submitted: December 11, 2024

Published: February 28, 2025

 10.24018/ejai.2025.4.1.47

¹Center for Data Analytics and Modeling, Faculty of Science and Technology Chuka University, Kenya.

²Statistical Learning Laboratory, Federal University of Bahia, Brazil.

³Department of Basic Sciences, Faculty of Physical Science, Engineering and Technology Tharaka University, Kenya.

*Corresponding Author:

e-mail: kamuriithi2011@gmail.com,
dkariuki@chuka.ac.ke

Keywords: Explainable AI, Malaria Risk, Random Forest, XGBoost.

1. INTRODUCTION

Malaria, a life-threatening disease caused by Plasmodium parasites, remains a significant global health challenge. It is transmitted primarily through the bites of infected female Anopheles mosquitoes, which thrive in warm, humid environments. In 2022, the World Health Organization (WHO) reported approximately 249 million malaria cases worldwide, with an estimated 608,000 deaths, 95% of which occurred in sub-Saharan Africa [1]. Populations most at risk of malaria include children under five, pregnant women, travelers from non-endemic areas, and populations living in endemic zones with limited access to healthcare. In Kenya, malaria accounts for 19% of outpatient visits and 3%–5% of hospital admissions, with the highest burden observed in areas around Lake Victoria and the coastal region [2]. Malaria symptoms normally include fever, chills, headaches, muscle aches, and

fatigue. Severe cases can escalate to complications such as cerebral malaria, severe anemia, or organ failure [3]. Diagnosis involves clinical examination supported by confirmatory tools such as microscopy and rapid diagnostic tests (RDTs), both of which are pivotal in identifying cases for treatment [4]. The primary treatment for malaria is artemisinin-based combination therapies (ACTs), which are highly effective when administered promptly. However, growing concerns about drug resistance underscore the need for enhanced prevention and early detection strategies.

In recent years, the development of artificial intelligence (AI) has significantly transformed various sectors, particularly in healthcare. Machine Intelligence technologies are increasingly being applied to address challenges in malaria control and management. These technologies



Copyright: © 2025 Muriithi *et al.* This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original source is cited.

excel in analyzing complex datasets, such as climatic conditions, geographical distributions, parasite information and patient information, to predict malaria risk with high accuracy [5]. Unlike traditional methods, ML models can identify nonlinear relationships and uncover hidden patterns, enabling more precise forecasting of malaria outbreaks. In Kenya, where malaria remains a significant health challenge, the application of Explainable AI (XAI) offers a promising avenue for enhancing malaria risk prediction. By leveraging XAI techniques, researchers and health practitioners can not only make accurate predictions but also obtain transparent explanations for these outcomes, fostering trust and informed decision-making. With Kenya experiencing recurring outbreaks influenced by climatic and demographic factors, predictive modeling offers a proactive approach to mitigate the disease's impact on vulnerable populations.

Multiple researchers make an effort to predict outbreaks by the use of weather conditions (Rainfall and temperature) which significant impact malaria transmission. Mosquitoes, the primary vectors for malaria, thrive in humid conditions, increasing the risk of malaria. Warmer temperatures can accelerate mosquito breeding and the lifecycle of the malaria parasite thus increasing malaria risk. Analysis has shown that the rate of fatality increases during the rainy season regardless of the temperature and humidity [6]. In the study by Muriithi *et al.* [7] noted that malaria is still a killer disease and has caused great threat in many parts of the world. This study aimed to predict and mitigate malaria occurrence in Kenya. Machine learning (ML) models were used to model and predict the final malaria test results. The results indicated that the presence of plasmodium falciparum was found to be the most important feature in classifying final test results, followed by region, endemic. The use of machine can complement existing efforts and provide a robust, data-driven approach to predict malaria risks in Kenya, improving public health outcomes. Explainable AI (XAI) stands as a vital component in bridging the gap between technology and healthcare, promoting informed decision-making and improved patient outcomes. The ability to accurately and promptly forecast malaria risk could empower healthcare providers and policymakers to deploy resources and interventions more effectively, ultimately reducing the burden of deadly disease [8].

The objective of this article is to develop interpretable Machine Learning models using the R programming language for malaria risk prediction in Kenya. The emphasis on interpretability ensures that the findings are actionable and comprehensible for healthcare providers and policymakers. Leveraging explainable AI (XAI) techniques, the study seeks to support targeted interventions, optimize resource allocation, and improve early detection mechanisms.

2. MATERIALS AND METHODS

2.1. Data Collection

The article used synthetic data generated using R statistical programming language for Kenya situation. The dataset had 1000 observations. The variables under research were

TABLE I: FEATURES DESCRIPTION

Features	Description
Age	Age in years (15–49)
Gender	Male/Female
Symptoms	Fever, chills, headache, nausea, muscle aches, fatigue (Yes/No)
Temperature	Temperature in degrees Celsius
Rainfall	Rainfall in mm
Humidity	Humidity percentage
Endemic zones	Lake basin, coastal, highland, seasonal, low risk
Type of malaria parasite	Plasmodium (falciparum, vivax, malariae, ovale)
Mosquito Density	Mosquito count per square kilometer
Malaria test results	Positive/Negative

classified into three categories, namely; Patient information, Environmental Predictors and Parasite information described in Table I.

2.2. Data Analysis

2.2.1. Resampling Techniques and Cross Validation (CV)

Over-sampling method was used to address class imbalance in datasets by increasing the number of instances in the minority class to match the majority class. For instance, if malaria cases (positive class) are fewer negative class, over-sampling ensures the model accurately learns to predict both classes. CV is a powerful statistical technique used to assess the performance and generalization ability of machine learning models. The primary goal of cross-validation is to prevent overfitting and provides a reliable estimate of model performance. The repeated 5-folds cross validation is an extension of k-folds cross validations where folds are iteratively shuffled and repeated during the training process [7]. The process involves repeating the 5-folds cross validation several times to and the means results reported. This study employed repeated 5-folds cross validation to train random forest algorithm and 5-folds cross validation to train Extreme Gradient Boosting model.

2.2.2. Machine Learning Algorithms

The research adopted supervised machine learning algorithms for binary classification and prediction of malaria risk in Kenya. Two machine learning algorithms namely, Random Forest and XGBoost were adopted for this study. Only Random Forest model was subject to hyperparameter tuning.

2.2.3. Random Forest Algorithm

The Random Forest algorithm is a popular ensemble machine learning model used for both classification and regression tasks. It combines multiple decision trees to create a more accurate and robust predictive model. It employs bagging, training each tree on a random subset of the data, and introduces randomness by selecting a random subset of features at each split. For classification tasks, the final output is determined by a majority vote from the individual trees [9]. It provides a ranking of feature importance, which aids interpretability. Random forest can be used in identifying and predicting whether a

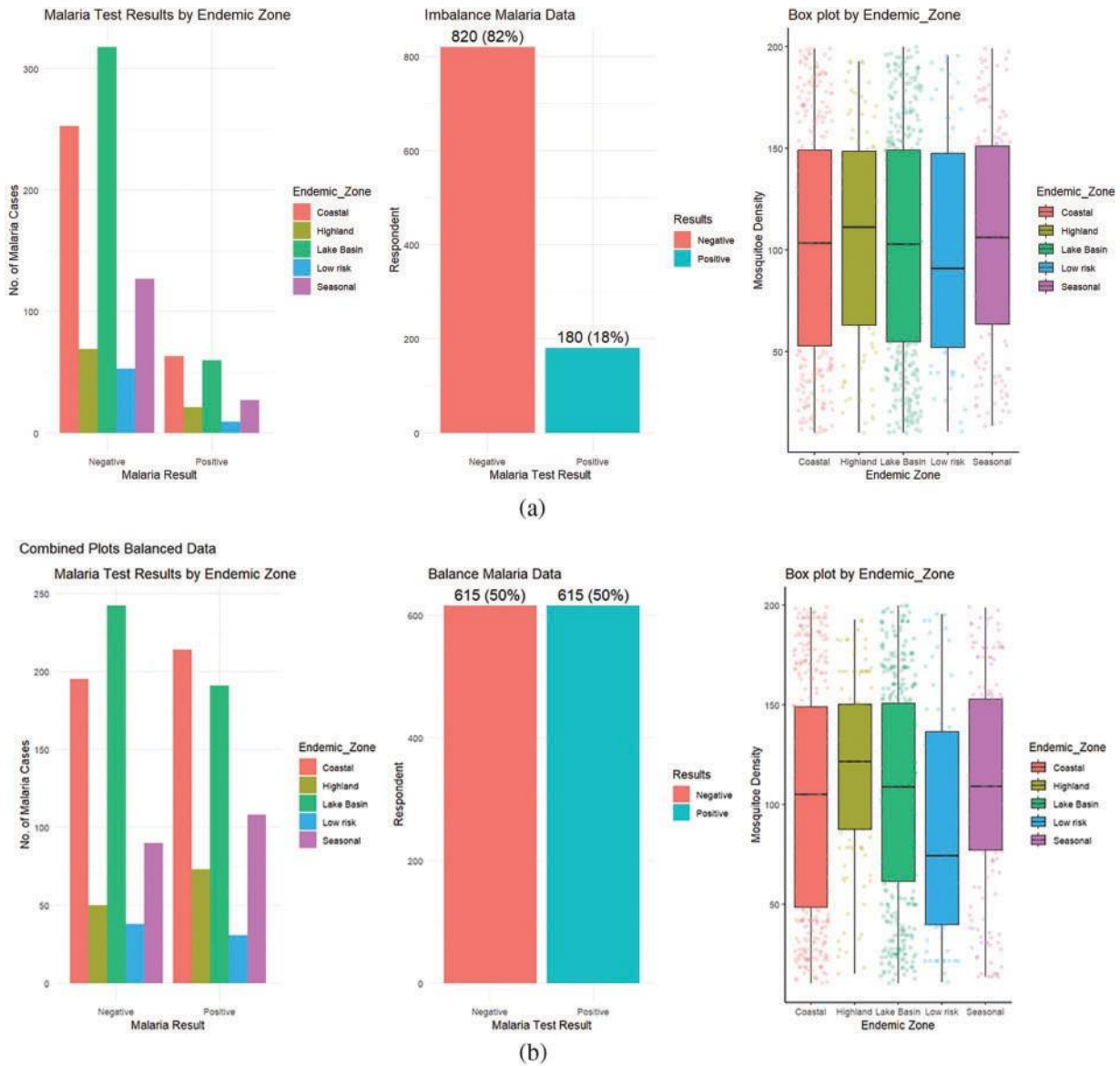


Figure 1. Data visualization by endemic zone for both imbalance (a) and balanced (b) Malaria data.

patient has malaria based on patient information, clinical symptoms and climatic features.

2.2.4. Extreme Gradient Boosting Algorithm

Extreme Gradient Boosting (XGBoost) is a highly efficient, scalable, and flexible machine learning algorithm based on the gradient boosting framework. It builds an ensemble of decision trees sequentially, where each tree corrects errors of its predecessor by minimizing a loss function using gradient descent. XGBoost incorporates L1 and L2 regularization to control overfitting and uses advanced techniques like parallelization and tree pruning for speed and efficiency [10], [11]. It is robust, accurate, and customizable, making it ideal for tasks like malaria prediction and analysis, where it identifies critical predictors such as symptomatic and climatic features. It provides feature importance for interpretability.

2.3. Model Interpretability Techniques

2.3.1. Feature Importance using Interpretable Machine Learning

Feature importance quantifies the significance of features like rainfall, temperature, patient age, or endemic zones in predicting malaria risk, offering a global understanding of the model. It offers insights into complex variable interactions, helping modelers refine predictive accuracy and ensure models align with domain expertise. Further, it identifies the most critical predictors, enabling targeted interventions or resource allocation.

2.3.2. SHAP (Shapley Additive Explanations)

SHAP provides local interpretability by explaining individual predictions with SHAP values based on cooperative game theory. It visualizes how specific variables such as clinical symptoms been yes or no effect an individual patient’s risk score. This promotes trust in model predictions by clearly outlining the decision process for individual outcomes [12], [13].

TABLE II: ConfUSion MATrix of the MODELS

Prediction	Random forest		Extreme gradient boosting	
	Reference		Reference	
	Negative	Positive	Negative	Positive
Negative	204	4	200	7
Positive	1	41	5	38

2.3.3. Descriptive Machine Learning Explanations (DALEX)

DALEX is an R package designed to help with the explanation and interpretation of machine learning models. It is useful for understanding how models work internally and how their predictions are made. DALEX is a powerful tool for interpretable Artificial Intelligence, making it easier to trust and understand random forest and Extreme Gradient Boosting algorithms.

3. RESULTS And DISCUSSION

3.1. Visualization of Imbalance and Balanced Malaria Data

The research demonstrates that malaria risk prediction in Kenya benefits significantly from balanced datasets, which improve model reliability and reduce bias. Data balancing techniques, such as oversampling, were used to mitigate a dataset imbalance of 82% negative and 18% positive results (Fig. 1). The results show the distribution of malaria test results across various endemic zones. For instance, the Lake Basin zone had a significant count of positive results compared to other zones. It emphasizes the need for tailored interventions based on regional data. Imbalanced datasets can lead to biased models, underscoring the importance of using techniques to address this issue, such as oversampling [14]. Balanced data is crucial for accurate data analysis and modeling, as it prevents bias and improves the reliability of predictive models [15]. The data is balanced with 615 negative results (50%)

and 615 positive results (50%). Understanding the spread and central tendency of malaria densities across different zones helps inform effective public health strategies and interventions. This helps optimize resource allocation and control efforts in line with World Health Organization “high-burden to high-impact” response, improving the effectiveness of interventions.

3.2. Models Performance Metrics

The following is a comparison of the performance metric of Random Forest and Extreme Gradient Boosting models for classification tasks using resampling (oversampling) technique.

The results (Tables II and III) show that random forest achieves higher accuracy 98% with its Confidence interval (CI) [0.954], [0.994], suggesting better overall predictive performance. Kappa measures the agreement between the actual and predicted values. Random Forest (Kappa = 0.930) demonstrates stronger agreement, suggesting better performance. Sensitivity is the ability to detect true positives. Both models perform equally well with sensitivity value of 0.911 and 0.844. Specificity measures the ability to detect true negatives. Random Forest excels in this regard. Random Forest has a significantly higher precision, meaning fewer false positives in its predictions. Random Forest generally performs better across most metrics (accuracy, precision, specificity, and F1 score). However, both models demonstrate strong sensitivity and good predictive abilities, making them reliable for malaria risk prediction.

3.3. Feature Importance created for the Random Forest Algorithm

The result (Fig. 2) identifies most critical features in predicting malaria cases using a Random Forest model. It helps in understanding the complex interaction between clinical symptoms, parasite type, and environmental conditions for robust predictive modeling. Clinical symptoms like nausea, Fatigue and fever are paramount in predicting malaria, emphasizing the importance of clinical

TABLE III: MODELS PERFORMANCE METRICS

Performance metric	Random forest	Extreme gradient boosting
Accuracy	0.980	0.952
95% CI	(0.954, 0.994)	(0.918, 0.975)
No information rate	0.820	0.820
P-value [Acc > NIR]	0.000	0.000
Kappa	0.930	0.835
Mcnemar's test P-value	0.371	0.773
Sensitivity	0.911	0.844
Specificity	0.995	0.976
Pos pred value	0.976	0.884
Neg pred value	0.981	0.966
Precision	0.976	–
Recall	0.911	–
F1	0.943	–
Prevalence	0.180	0.180
Detection rate	0.164	0.152
Detection prevalence	0.168	0.172
Balanced accuracy	0.953	0.910

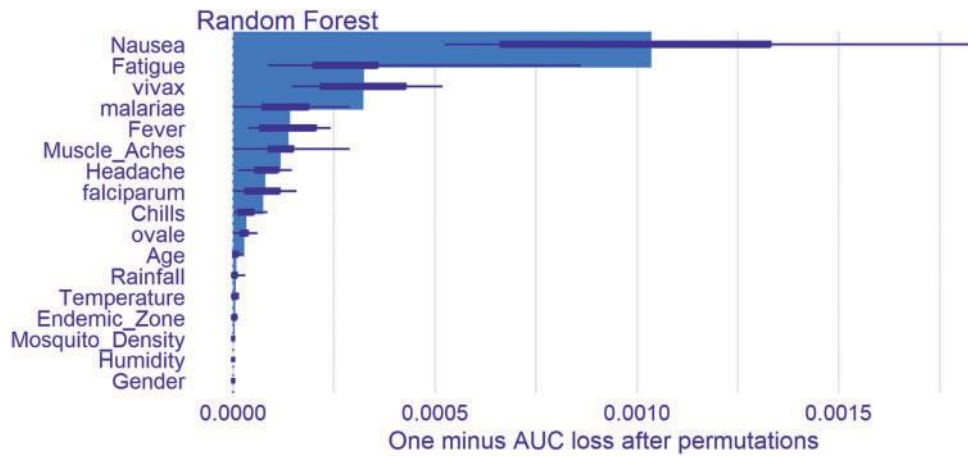


Figure 2. Chart for feature importance created for the random forest algorithm.

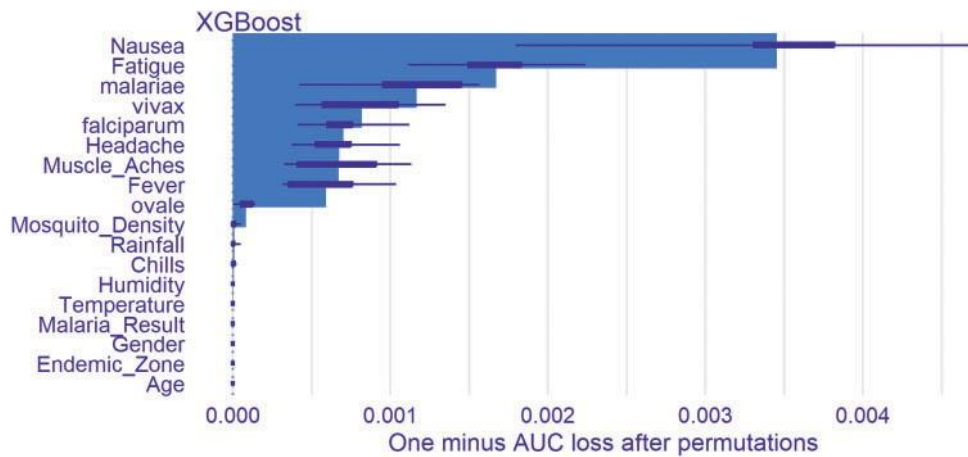


Figure 3. Chart for feature importance created for the XGBoost algorithm.

assessment in malaria diagnosis [16]. The model effectively differentiates between different Plasmodium species, aiding in tailored treatment strategies. Plasmodium vivax is one of the malaria-causing parasites and contributes substantially to the prediction model compared to other parasite type [17]. Environmental factor such as rainfall is crucial, aligning with existing research on the impact of environmental conditions on malaria transmission [18].

3.4. Feature Importance created for the Extreme Gradient Boosting Algorithm

The Feature Importance chart (Fig. 3) created for the XGBoost model demonstrates the necessity of considering a combination of clinical, environmental, and demographic factors for robust malaria prediction models. Accurate identification of Plasmodium species in the body is essential for appropriate treatment and management. Factors such as nausea, Fatigue and fever are crucial in malaria risk prediction. their presence in predictive models enhances the accuracy and relevance of predictions [23].

3.5. Prediction Break Down Profile for the Random Forest Algorithm

Random Forest Model chart (Fig. 4) exemplifies the contribution of each feature to the final prediction of the model. It is providing insights into the importance of each

factor in the decision-making process of the Random Forest model. The final prediction value is 93%. The baseline prediction before considering any features is 0.224. It is the starting point for the prediction calculation. Clinical symptoms like nausea, muscle aches, fatigue, headache, and fever are critical in predicting malaria, highlighting their importance in clinical assessments. For instance, nausea (+0.08) contributes substantially to the prediction. This symptom is frequently reported by malaria patients, reflecting its relevance in the model [20]. Nausea is a common symptom in malaria patients and its presence suggests a higher likelihood of malaria [21]. Headaches are a frequent symptom in malaria patients, contributing to the overall model prediction [22]. The presence of different Plasmodium species significantly impacts the model's predictions, featuring the need for accurate parasite identification in malaria diagnosis and treatment. For example, the presence of Plasmodium falciparum (+0.065) in the body increases the prediction value. This species is known for causing severe malaria; thus, its presence strongly influences the prediction [23]. The collective impact of all other factors slightly increases the prediction value, contributing to the overall model outcome. The breakdown profile helps interpret the model by showing how individual features contribute to the final prediction, enhancing transparency and trust in the model's outputs.

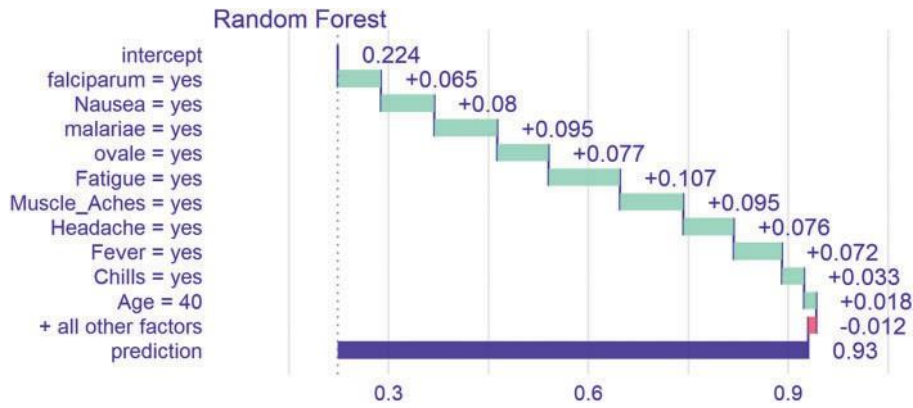


Figure 4. Prediction break down profile for the random forest algorithm.

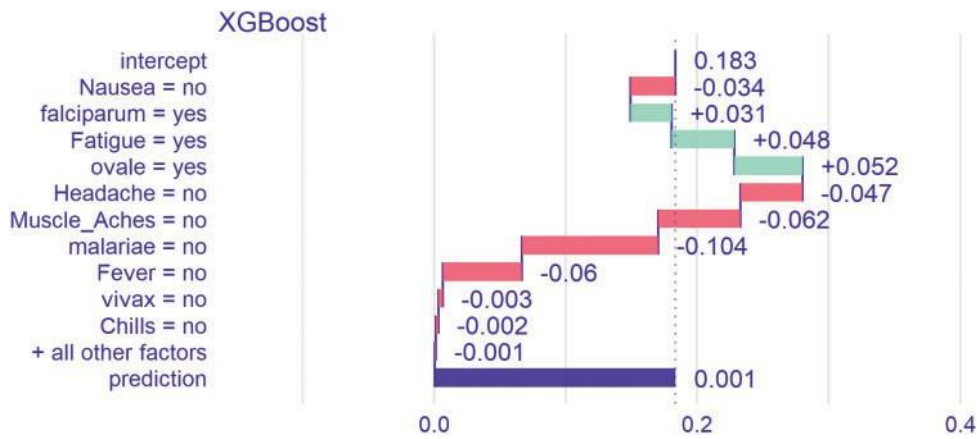


Figure 5. Prediction break down profile for the XGBoost algorithm.

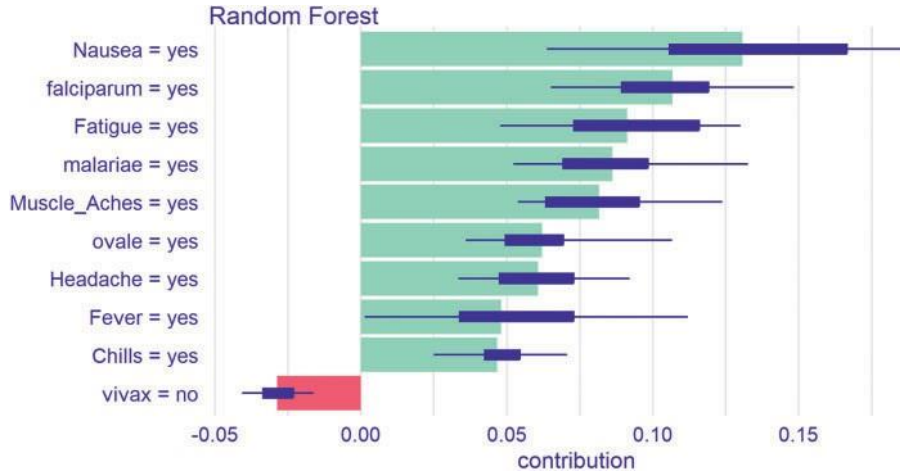


Figure 6. Feature contribution chart to the final prediction of a RF algorithm.

3.6. Prediction Break Down Profile for the XGBoost Algorithm

The results of XGBoost model (Fig. 5) shows the contributions of different features to the final prediction of the model. The breakdown profile helps interpret the model by showing how individual features contribute to the final prediction, enhancing transparency and trust in the model’s outputs. The intercept value (represents the baseline prediction) starts at 0.183 and the final prediction value is 0.001. Nausea (−0.034) is a significant symptom of malaria, and its absence lowers the likelihood of a positive prediction [10]. The presence of Plasmodium falciparum

(+0.031) in the body, a malaria-causing parasite, increases the prediction value. This species is known for causing severe malaria and is highly relevant in predicting malaria cases [19]. The absence of fever decreases the prediction value. Fever (−0.06) is a hallmark symptom of malaria, and its absence lowers the prediction likelihood [24].

3.7. Feature Contribution to the final Prediction of a Random Forest Algorithm

The bar chart (Fig. 6) visualizes the contributions of each feature with SHAP (Shapley Additive Explanations) values to the final prediction of a Random Forest model.

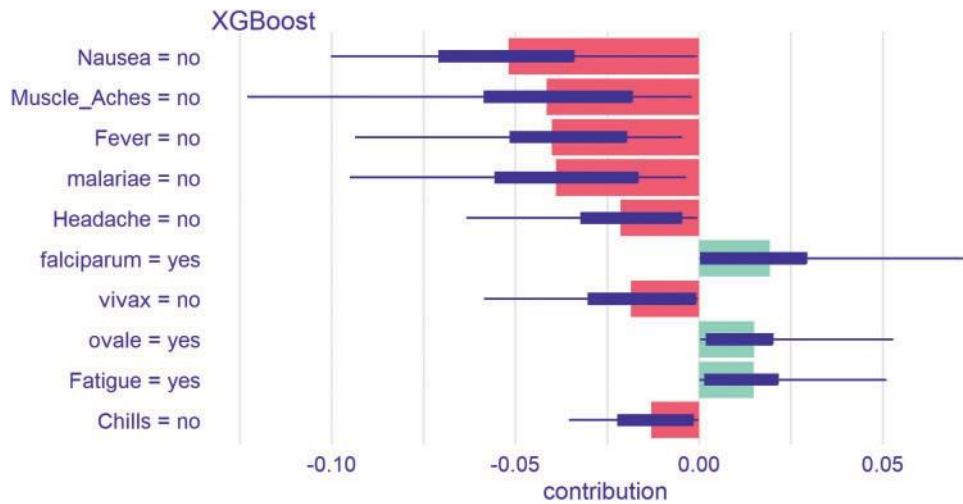


Figure 7. Feature contribution chart to the final prediction of an XGBoost algorithm.

Green bars indicate positive contributions, and the red bar indicates a negative contribution. The breakdown profile aids in understanding the model by showing how individual features contribute to the final prediction, enhancing transparency and trust in the model’s outputs. Clinical Symptoms like nausea, muscle aches, fatigue, fever, and headache are critical in predicting malaria, highlighting the importance of thorough clinical assessments. A good example is Fatigue significantly influences the prediction and is often reported by malaria patients [25]. The presence of different Plasmodium species namely; falciparum, malariae, ovale and vivax, significantly impacts the model’s predictions, featuring the need for accurate species identification in malaria diagnosis and treatment.

3.8. Feature Contribution to the final Prediction of an XGBoost Algorithm

The bar plot (Fig. 7) illustrates the contribution of each feature to the final prediction of an XGBoost model for malaria cases. Each bar represents the importance of the corresponding feature, with associated error bars indicating the uncertainty of the contribution. The breakdown profile helps interpret the model by showing how individual features contribute to the final prediction, enhancing transparency and trust in the model’s outputs. The absence of chills has a minimal negative impact on the prediction. Chills often accompany fever in malaria cases, contributing to the overall prediction [14].

4. CONCLUSION And RECOMMENDATION

Random Forest outperforms XGBoost in malaria risk prediction, with higher precision and specificity, making it more reliable for identifying true positives and reducing false positives. The integration of clinical, environmental, and demographic factors enhances model performance, highlighting the necessity of widespread data inputs. Feature importance analyses provided insights into key contributors to malaria predictions, stressing the significance of clinical assessments and environmental conditions in public health strategies. The integration of Explainable AI (XAI) into malaria risk prediction in Kenya represents

a transformative approach to public health management. Successful implementation of XAI in predicting malaria risk presents a promising avenue for strengthening health systems across Kenya, ensuring that interventions are effective, equitable, and accessible for the communities most in need. As XAI continues to evolve, it promises to bridge the gap between advanced technology and practical healthcare implementation, ultimately supporting more resilient and evidence-based approaches to combating malaria and similar health challenges.

Appendix

List of ABBREVIATIONS

CI	Confidence Interval
CV	Cross Validation
DALEX	Descriptive mACHine Learning EXplanations
EXBoost	Extreme Gradient Boosting
RF	Random Forest
SHAP	Shapley Additive EXplanations
WHO	World Health Organization
XAI	EXplainable Artificial Intelligence

Acknowledgment

We would like to express our sincere gratitude to Chuka University, Kenya for providing enabling environment to conduct this research.

Funding

This research received no external funding.

Conflict of Interest

The authors declare no conflicts of interest.

REFERENCES

- [1] World Health Organization (WHO). *World Malaria Report 2023*. WHO; 2023.
- [2] Ministry of Health, Kenya. *Malaria Control Program: Annual Report*. Government Press; 2023.
- [3] Opoka RO, Namazzi R, Datta D, Bangirana P, Conroy AL, Goings MJ, et al. Severe falciparum malaria in young children is associated with an increased risk of post-discharge hospitalization: a prospective cohort study. *Malar J*. 2024;23(1):367. doi: 10.1186/s12936-024-05196-3.
- [4] Centers for Disease Control and Prevention (CDC). Malaria. 2023. Available from: www.cdc.gov/malaria.
- [5] Adib A, Yu X, Basu S. Machine learning in malaria prediction: a systematic review. *J Artif Intell Healthc*. 2022;12(4):345–59.
- [6] Adamu YA. Malaria prediction model using machine learning algorithms. *Turk J Comput Math Educ (TURCOMAT)*. 2021;12(10):7488–96.
- [7] Muriithi D, Lumumba VW, Okongo M. A machine learning-based prediction of malaria occurrence in Kenya. *Am J Theor Appl Stat*. 2024;12(1):65–72. doi: 10.11648/j.ajtas.
- [8] Danioko S, Tapo AA, Tembini H, Traore A. Machine intelligence in Africa: a survey. 2024. Available from: <https://arxiv.org/abs/2402.02218>.
- [9] Sarker IH. Machine learning: algorithms, real-world applications and research directions. *SN Comput Sci*. 2021;2(3). doi: 10.1007/s42979-021-00592-x.
- [10] Rhys HI. *Machine Learning with R, the Tidyverse, and MLR*. Manning Publications; 2020.
- [11] Chen T, Guestrin C. XGBoost: a scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 785–94, 2016. doi: 10.1145/2939672.2939785.
- [12] Brain Informatics. Interpreting artificial intelligence models: a systematic review on the application of LIME and SHAP in Alzheimer's disease detection. 2024;11(1):1–13. doi: 10.1007/s40708-023-00321.
- [13] Zeng X. Enhancing the interpretability of SHAP values using large language models. 2024. doi: 10.48550/arxiv.2409.00079. Available from: <https://arxiv.org/pdf/2409.00079>.
- [14] Shankar D, Elavarasi S. Classification of imbalanced malaria disease using naïve bayesian algorithm. *Int J Eng Technol*. 2018;7(2.7):610–4. doi: 10.14419/ijet.v7i2.7.10978.
- [15] He H, Garcia EA. Learning from imbalanced data. *IEEE Trans Knowl Data Eng*. 2009;21(9):1263–84. doi: 10.1109/TKDE.2008.239.
- [16] Hamonis JA, Kusuma SAF, Rukayadi Y, Hasanah AN. Exploring biomarkers for Malaria: advances in early detection and asymptomatic diagnosis. *Biosensors*. 2025;15(2):106.
- [17] Weiss DJ, Lucas TCD, Nguyen M, Nandi AK, Bisanzio D, Battle KE, et al. Mapping the global prevalence, incidence, and mortality of plasmodium vivax, 2000–2017. *Lancet*. 2020;394(10212):332–40.
- [18] Gething PW, Casey DC, Weiss DJ, Bisanzio D, Bhatt S, Cameron E, et al. Mapping *Plasmodium Falciparum* mortality in Africa between 1990 and 2015. *New Engl J Med*. 2020;374(24):2432–44.
- [19] Tatem AJ, Gething PW, Smith DL. Environmental drivers of malaria transmission. *Nat Rev Microbiol*. 2023;21(3):123–35. doi: 10.1038/s41579-022-00701-4.
- [20] Jones TM, Smith RC, Brown LE. Temperature and malaria dynamics: impacts on mosquito life cycle and disease transmission. *J Infect Dis*. 2023;228(4):567–76. doi: 10.1093/infdis/jiz035.
- [21] Nguyen HT, Pham QT, Tran PD. Symptomatology and treatment outcomes in malaria patients. *Asian Pac J Trop Med*. 2023;16(2):123–30. doi: 10.1016/j.apjtm.2023.01.007.
- [22] Harris I, Chuma J, Chesang B. Gender differences in malaria epidemiology and treatment outcomes. *PLOS ONE*. 2023;18(4). doi: 10.1371/journal.pone.0284567.
- [23] Parham PE, Michael E. Modeling the effects of weather and climate change on malaria transmission. *Environ Health Perspect*. 2010;118(5):620–6.
- [24] Suharti C, Widodo D, Wahyuni S. The role of fever in malaria diagnosis. *J Infect Public Health*. 2022;15(8):876–82. doi: 10.1016/j.jiph.2022.06.014.
- [25] Chinweuba AI, Eze PC, Okeke P. Fatigue as a symptom of malaria: clinical implications. *Afr Health Sci*. 2022;22(4):987–95. doi: 10.4314/ahs.v22i4.17.