[redacted]

American Journal of Theoretical and Applied Statistics

2024, Vol. 13, No. 4, pp. 65-72

https://doi.org/10.11648/j.ajtas.20241304.11

Research Article

A Machine Learning-Based Prediction of Malaria Occurrence in Kenya

Dennis Muriithi* , Victor Wandera Lumumba

, Mark Okongo

Gmail: okongo@chuka.ac.ke

Center for Data Analytics and Modelling, Faculty of Science and Technology, Chuka University, Chuka, Kenya **Abstract**

For many years' malaria has been a health public concern in Kenya as well as many parts of Africa and other parts of the world.

The purpose of this study is to develop and evaluate a supervised machine learning model to predict malaria occurrence (final malaria test results) in Kenya. The study investigated twelve predictor variables on the outcome variable (malaria test results), where five machine learning models namely; k-nearest neighbors, support vector machines, random forest, tree bagging, and boosting, were estimated. During the model evaluation, random forest emerged as the best overall model in the classification and prediction of final malaria test results. The model attained a higher classification accuracy of 97.33%, sensitivity of 71.1%, specificity of 98.4%, balanced accuracy of 84.7% and an area under the curve of 98.3%. From the final model, the presence of plasmodium falciparum emerged most important feature, followed by region, endemic zone and anemic level. The feature with the least importance in predicting final malaria test results was having mosquito nets. In conclusion, employing Machine learning algorithms enhances early detection, optimizing resource allocation for interventions, and ultimately reducing the incidence and impact of malaria in the Kenya. The study recommends allocation of resources and funds to areas with the presence of plasmodium falciparum, region susceptible to malaria, endemic zones and anemic prone areas.

Keywords

Machine Learning, Accuracy, Sensitivity, Specificity, Feature, Balance Accuracy, Malaria **1. Introduction**

Malaria is a killer disease and has caused great threat in the other hand, life threatening signs and symptoms include many regions especially in the malaria tropical regions and confusion, seizures, jaundice, dark urine, and difficulty endemic zones [19]. However, the disease is considered breathing among others. The threat from malaria infection deadly but curable. Unlike other diseases, malaria is caused varies significantly from one group to another. In a report by and spread by female Anopheles mosquitos which carries World Health Organization 2022, children under five (5) plasmodium parasite and is not transmitted from one person to years, pregnant women, and travelers are at great risk of the another [12]. Malaria infection is accompanied by quite a threat caused by this killer disease. It is important to note that number of signs and symptoms which can be regarded as mild malaria infection is not caused by a single type of malaria and some regarded as deadly [15]. Some of the mild malaria parasite. There exist five type of mosquito parasite from the signs and symptoms include fever, headache and chills. On female anopheles' mosquitos that causes malaria and the two

*Corresponding author:

Received: 20 July 2024; **Accepted:** 9 August 2024; **Published:** 20 August 2024

American Journal of Theoretical and Applied Statistics

http://www.sciencepg.com/journal/ajtas of them are plasmodium *falciparum* and plasmodium *vivax* ia-related mortality rate for children under five years between

[14]. These species of malaria parasites exist in various re-2003 and 2022 reduced from 11.5% to 4.1%, indicating 0.389%

gions, however, the most prevalent type of malaria parasite decrease in malaria-related mortality rate among children that exists in most parts of Africa is the plasmodium *falci-under five years. The decrease in mortality in this group was parum [11]. The species is the most threatening malaria spe-made possible by allocating more funds and resource to area cies. Plasmodium vivax is most prevalent in other countries in such as high endemic zones and lake endemic zones in ac-Africa outside the sub-Saharan parts of Africa. The other three cordance with the PMI initiatives.*

species of malaria parasite are plasmodium malariae, plas-It is worth acknowledging that much has been done to ad-modium ovale and plasmodium knowlesi.

dress the malaria infection cases in Kenya as well as many In their 2022 report, the World Health Organization (WHO) parts of world, however, this paper seeks to compliment what reported that there were 249 million malaria cases with ap-has been done by developing machine learning predictive proximately 608,000 malaria-related deaths in 2022 as com-models to model and predict malaria occurrence in Kenya.

pared to 610,000 malaria-related deaths in 2021. Despite the Application of machine learning in binary and mul-decrease in malaria-related deaths in 2022, the disease in still ti-classification is relevant due to the ability of the algorithms a life-threatening and requires continuous and proactive to analyze vast amount of data to uncover hidden insights and measures to prevent its resurgence and manage its transmis-patterns that could not be uncovered by the traditional meth-sion effectively. These statistics are reported from 85 coun-ods [6]. Despite the reports showing a reduction in the matries. Out of the number of malaria-related deaths reported laria-related mortality rate, malaria infection is still a public globally, nearly half the number of deaths is reported from health concern in Kenya. Application of machine learning in four African countries which include Nigeria, Uganda, this study made use of various factors including demographic Democratic Republic of Congo (DRC) and Mozambique. In factors, environmental factors and health related factors to their study, [3] reported that higher share of malaria cases is accurately predict malaria occurrences. The accuracy in the disproportionately higher in African countries as compared to prediction of malaria cases in this study was made possible any other country. Besides, approximately, 95% of death in due to the ability of machine learning models to analyze and Africa are malaria-related deaths which is close to 580,000

derive insights from both linear and non-linear relationship deaths [16]. In the general population, children under five (5) between features. This ability in deriving insights from com-years of age are at a great risk of malaria infection and ma-plex data set is vital in developing an intervention program to laria-related death. In their report, WHO reports that 80% of address malaria-related threats especially among children the deaths among children *under five years were found to be under five years, pregnant women and travelers. Further, malaria related.*

application of machine learning algorithms in this study will Kenya as one of the African countries in sub-Saharan part help ensure that prediction of malaria case is accurate over of Africa, faces the same threat from malaria infection just time as more data get incorporated into the model which like any other African tropical countries [10]. Many countries increase the performance of the ML models. Therefore, the lying within 350 S and 350 N are likely to fall in the tropical integration of machine learning in this study aimed to com-region [5]. Since the equator, tropic of Cancer, tropic of plement existing efforts and provide a robust, data-driven Capricorn nearly runs through the middle of Africa, it makes approach to predict and mitigate malaria occurrence, im-Africa the most tropical continent resulting to higher suscep-proving public health outcomes in Kenya. The purpose of this tibility to malaria infection. The tropical climate of Africa, study is thus to develop and evaluate machine learning models characterized by warm temperatures higher humidity and to predict malaria occurrence in Kenya, with the objectives of sufficient rainfall, creates favorable conditions for the enhancing early detection, optimizing resource allocation for breeding of mosquitoes, which are the primary vectors for interventions, and ultimately reducing the incidence and im-malaria, making the continent highly susceptible to malaria pact of malaria in the Country.

infection [8]. Equator passing through Kenya places the country in the tropical region with warm temperature and higher humidity, conditions ideal for the breeding and sur-2. Methodology

vival of anopheles' mosquitoes responsible for the transmis-sion of malaria [7]. These climatic conditions make the ***2.1. Data Collection***

country susceptible to malaria infection, a great public health concern. Several initiatives have been put in places in at-The data used in this study was obtained from the Kenya tempts to reduce malaria infection cases and deaths, however, National Data Archive (KeNADA) website using the link the reduction in the number of cases and deaths has not been https://statistics.knbs.or.ke/nada/index.php/catalog/111/relate significant [20]. In Kenya, malaria infection cases and deaths d-materials. The data used was well documented, accurate and are still relatively higher and a robust action is needed to relevant in addressing the research objectives in this study.

address and mitigate the infection. In a report by US Presi-The dataset had 31,302 observations with 223 variables. Upon dential Malaria Initiative (PMI) 2022, the number of malar-cleaning the data to remain with the most relevant information, we remained with thirteen variables comprising twelve pre-66

 dictors including region, endemic zones, anemic level, num-distance [17].

ber of mosquito bed nets, mother's educational level, presence of various plasmodium species among other variables. The �(�,�-, �,�-) = √∑� (�,�- − �,�-)2

predictors were all categorical and coded appropriately. The �=1

(7)

�

�

outcome variable in this study is the final malaria test results showing either positive or negative, indicating that an indi-The aggregation of neighbors' output is found as shown; vidual is infected or not infected, respectively [1].

�̂ = mode(�� for � ∈ N�) (8) ***2.2. Data Analysis***

The predicted class �̂ for the test instance, � is the class The paper adopted supervised machine learning algorithms that appears most frequently among the � selected neighbors: for binary classification and prediction of malaria occurrences arg ���

in Kenya. Five machine learning algorithms were adopted, �̂ =

(∑�

�∈� �(�� = �)) (9)

�

c∈C

namely; Support Vector Machines (SVM), K-Nearest Neighbors (K-NN), Random Forest, Tree Bagging, and 3) Random Forest

Boosting.

This algorithm is an ensemble that uses the majority voting 1) Support Vector Machines

as indicated by the formula below to increase the accuracy and Training the SVM model involves solving two optimiza-reduce overfitting [13];

tion problems in primal and dual [2]. The primal and the dual optimization problem is expressed as shown below; �̂ = Ar

ﹸ g

�

□□

∑□□=1 □ (□□(□) = □)

Primal form;

(10)

□∈□

□□□1 ||□||2 + □ ∑□ □ (1) *Letting T be the number of trees in the developed random* □,□,□2

□=1 □

forest model, and □̂□(□) *be the predictor of the tth tree, for instance, x. The final prediction will be given by; Solving the primal optimization problem is subject to the following conditions;*

□̂(□) = □□□□(*□̂

□

□(□)+□=1) (11)

□□(□ · □□ + □) ≥ 1 − □□, ≥ 0, □ = 1, … , □ (2) *4) Tree Bagging*

Give the dataset with n sample (□□,□□), *the algorithm The dual form;*

creates B bootstrap samples 1, 2 … □ . *The algorithm*

^

□□□

trains the decision tree □□ *on each bootstrap sample* □ *For the*

.∑□ □

□

□=1 □ − 1 ∑

□

/ (3)

2

□,□=1 □□□□□□□□(□□□□

new input X, the prediction is given as shown in the equation □

12.

The solution to the optimization problem above is subject to the following;

□□□□□□□□□□□□□□□: □̂ = □□□□□(□̂1, □̂2, … , □̂□) (12)

∑□ □

5) Boosting

□=1 □□□ = 0 , 0 ≤ □□ ≤ □, □ = 1, … , □ (4) Ada boost model is an ensemble that combines multiple The final model is expressed in terms of support vector weak learners from multiple decisions [4]. In this method, expressed as follows;

the boosting algorithm assigns equal weight to all training samples. The □□ is given by 1 where i =1, 2, …, N where N is □(□) = ∑□ □

□=1 □□□ □(□□, □□) + □ (5) the number of training samples. The model is developed using several weak learners m = 1,2, 3,…, M. The weaker For the new input feature x (test set), the model predicts learner □□(□) is developed from the weighted training sam-the class label (Positive or Negative) using the sign ples. Each of the weak learners developed is accompanied by of □(□) as give in the equation 6 below weighted error rate and the learner's weight as given by the equation 13 and 14.

□□□□□□□□□ □□□□□ = □□□□(∑□ □

□=1 □□□□(□□, □□) + □) (6)

□

(□)

□□□□□□□□ □□□□□ □□□□ = □

□(□

□=1

□≠h □(□□))

2) K-Nearest Neighbors

□ = ∑

□□

(13)

∑□ □(□)

□=1 □

The concept of the K-NN algorithm is built behind the idea of commonalities and neighbors' distance around the □□□□□□□′□ □□□□□□□ = □□ = 1 □□ .1−□□/ (14) response variable's target class known as k-Nearest Neigh-2

□□

bors determined by the distance metric known as Euclidean 67

http://www.sciencepg.com/journal/ajtas Each learner's weight is updated iteratively as shown in training set was used to estimate the model and the test set equation 15;

used to evaluate the model's performance.

□□□□□□□□ □□□□□□□□′□□□□□□□□ = □(□+1) =

□

3. Results and Discussions

□(□)exp (−□

□

□□□□□(□□)) (15)

3.1. Descriptive Statistics

The final step of training the ada boosting model is the weight normalization

(□+1)

Table 1. *Distribution of Malaria Test Results.*

□□□□□□□ □□□□□□□□□□□□□□ = □(□+1) = □□

(16)

∑ (+1)

=1

Malaria Test Results

N = 3,280

= (∑

=1 ()) (17)

Negative

3149 (96%)

Steps in Machine Learning Modeling Positive

131(4%)

Machine learning modeling takes quite a number of steps.

The initial step in ML model development was the problem formulation followed data acquisition and then data prepro-A total of 3280 households were enrolled in the study.

cessing. In the preprocessing step, the data was cleaned to Majority of households, 96.0% (n = 3,149) tested negative and remain with the most relevant information. Data partitioning 4.0% (n = 131) positive of malaria as shown in Table 1 above.

was to obtain the training set and testing set, where the This is a case of imbalance data set.

Table 2. *Two way Table Showing the Distribution of Malaria Test Results Across Various Factors.*

Characteristic

Negative, N = 3,149 1

95% CI 2

Positive, N = 131 1

95% CI 2

p-value 3

Endemic Zones

Highland Epidemic

540 (17%)

16%, 19%

4 (3.1%)

0.98%, 8.1%

Lake Endemic

1,088 (35%)

33%, 36%

107 (82%)

74%, 88%

Coastal Endemic

353 (11%)

10%, 12%

14 (11%)

6.2%, 18%

<0.001

Seasonal

743 (24%)

22%, 25%

6 (4.6%)

1.9%, 10%

Low Risk

425 (13%)

12%, 15%

0 (0%)

0.00%, 3.6%

Number of Children Slept Under Net Last Night None

1,202 (38%)

36%, 40%

40 (31%)

23%, 39%

One

1,269 (40%)

39%, 42%

70 (53%)

45%, 62%

Two

585 (19%)

17%, 20%

17 (13%)

8.0%, 20%

0.008

Three

82 (2.6%)

2.1%, 3.2%

2 (1.5%)

0.26%, 6.0%

Four

11 (0.3%)

0.18%, 0.64%

2 (1.5%)

0.26%, 6.0%

Anemia Level

Severe

69 (2.2%)

1.7%, 2.8%

8 (6.1%)

2.9%, 12%

Moderate

774 (25%)

23%, 26%

66 (50%)

42%, 59%

<0.001

Mild

749 (24%)

22%, 25%

32 (24%)

18%, 33%

Not anemic

1,557 (49%)

48%, 51%

25 (19%)

13%, 27%

Mother's Highest Educational Level

No education

525 (17%)

15%, 18%

10 (7.6%)

3.9%, 14%

<0.001

Primary

1,428 (45%)

44%, 47%

89 (68%)

59%, 76%

68

 Characteristic

Negative, N = 3,149 1

95% CI 2

Positive, N = 131 1

95% CI 2

p-value 3

Secondary

866 (28%)

26%, 29%

26 (20%)

14%, 28%

Higher

330 (10%)

9.4%, 12%

6 (4.6%)

1.9%, 10%

Table 2 shows two-way distribution of malaria test cases 3, random Forest emerges as the best overall model. The across endemic zone, number of children who slept under net model achieves the highest sensitivity (0.711) and a strong last night, anemic level, and mother highest education level.

specificity (0.984), indicating that the model effectively From the results, there is a statistically significant association identifies both positive and negative cases. Its precision between the factors identified in the results and the malaria (0.643) and F1-Score (0.675) reflect a good balance between test results.

precision and recall, ensuring reliable positive predictions.

Moreover, Random Forest has the highest balanced accuracy ***3.2. Model Estimation and Validation*** *(0.847), demonstrating its superior capability in handling imbalanced datasets compared to other models. While Model estimation in this paper was done with ten folds boosting also performs well with high specificity (0.987) and cross validation repeated five times, using repeated cross precision (0.657), its slightly lower sensitivity (0.605) and validation. Besides, the class function adopted in this study balanced accuracy (0.796) make it less optimal than Random was the two-class summary since the outcome variable is a Forest. Hence, Random Forest stands out as the most robust binary variable with (0=Negative, 1 = Positive). The results of and balanced model for this classification task. The results are ML models are reported in Table 3.*

in line with a study by [9] who found out that random forest emerged the best overall model in malaria prediction after ***3.3. Models Evaluation***

applying SMOTE. However, the results in this study are in-consistent with what [18] found out, who in their study, Naï ve

3.3.1. Models Performance Metrics

Bayes outperformed kNN, SVM and logistic regression in predicting malaria outbreak.

Considering the given performance metrics results in Table

Table 3. *Model's Performance Evaluation.*

Performance Metrics

Support Vector Machines

K Nearest Neighbors

Random Forest

Tree Bagging

Boosting

Sensitivity

0.447

0.053

0.711

0.711

0.605

Specificity

0.985

0.999

0.984

0.974

0.987

Precision

0.548

0.667

0.643

0.529

0.657

F1-Score

0.493

0.098

0.675

0.607

0.63

Balanced Accuracy

0.716

0.526

0.847

0.842

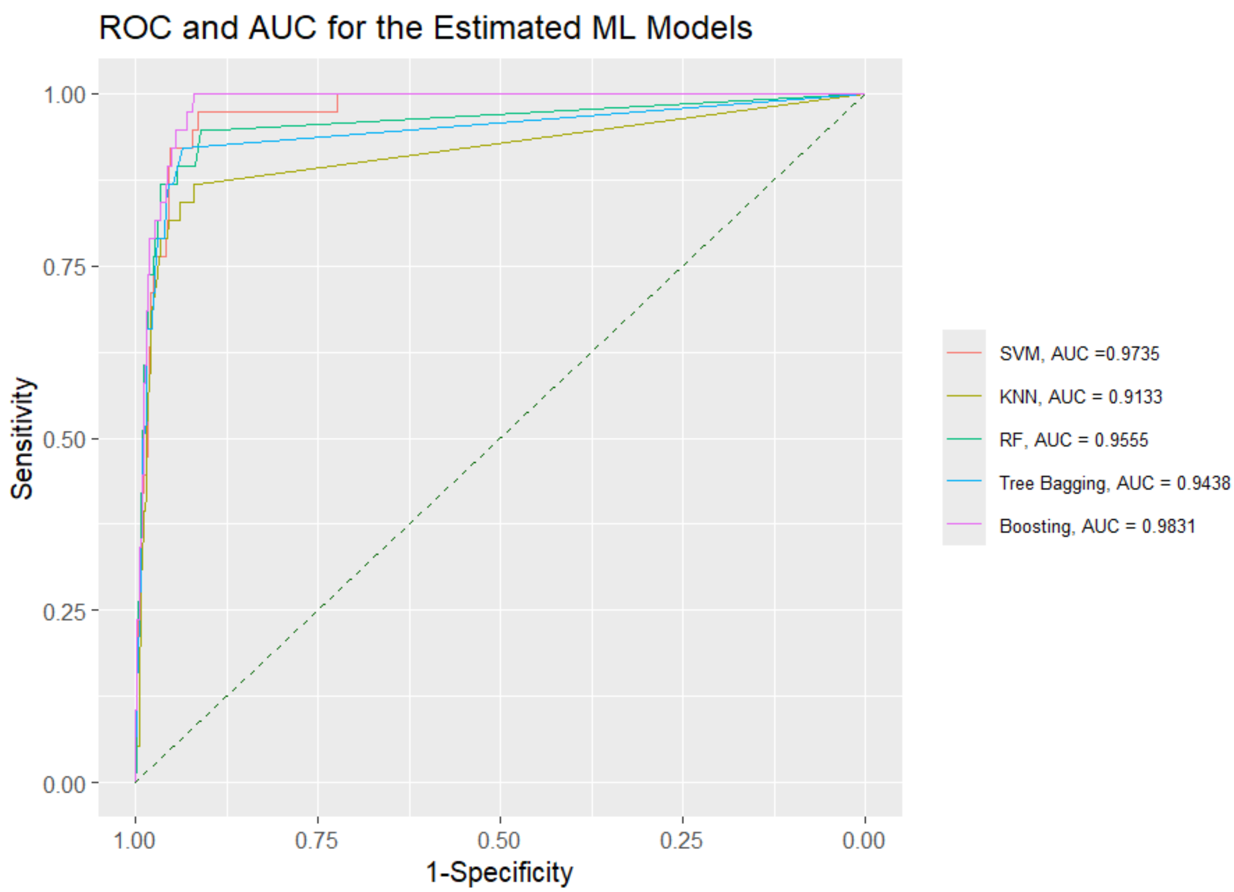
0.796

Considering the performance metrics in Table 3 together ***3.3.2. Receiver Operating Characteristic (ROC) and*** *with the results in Figure 1, random forest demonstrates a* ***Area under the Curve (AUC)***

higher ability in classifying the positive cases and the overall classification ability. While boosting has the highest AUC and Received operating curve (ROC) and area under the curve strong performance metrics, random forest's superior sensi-

(AUC) are important performance metrics in machine learn-tivity and balanced accuracy make it the most robust and ing especially for binary. The ROC and AUC estimated from reliable model for this classification task.

our ML models in this study are reported in Figure 1.


ROC and AUC for the Estimated ML Models
Sensitivity
1-Specificity
1.00
0.75
0.50
0.25
0.00
SVM, AUC =0.9735
KNN, AUC = 0.9133
RF, AUC = 0.9555
Tree Bagging, AUC = 0.9438
Boosting, AUC = 0.9831

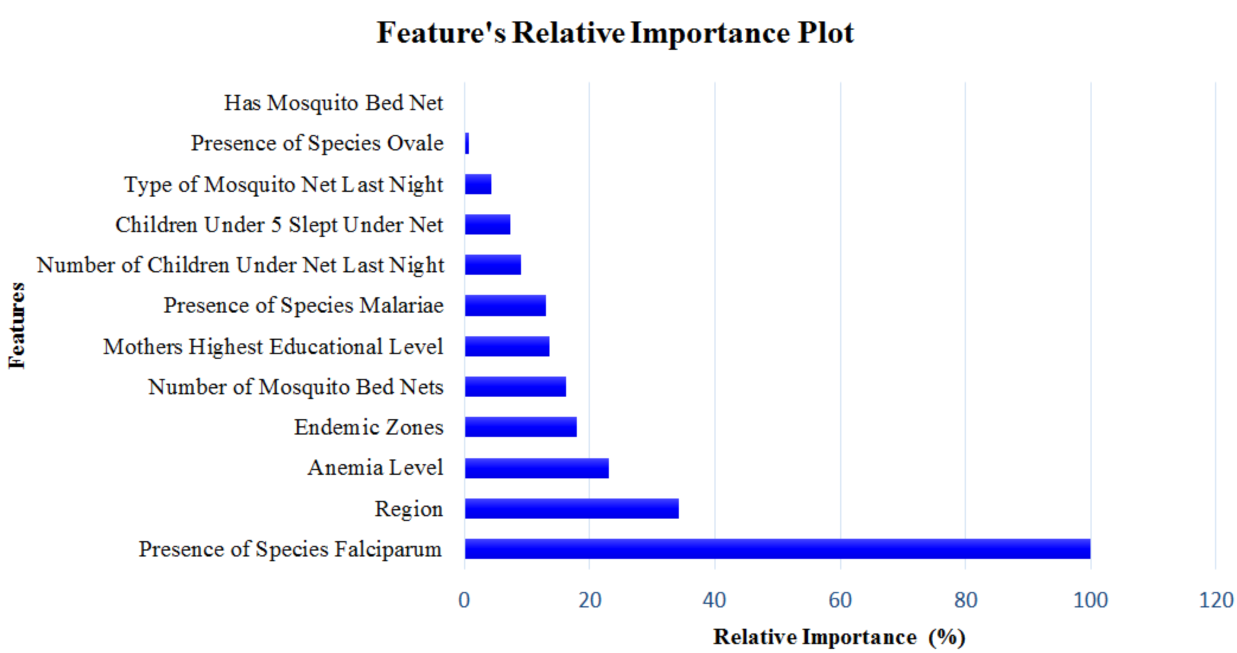

Feature's Relative Importance Plot
Has Mosquito Bed Net
Presence of Species Ovale
Type of Mosquito Net Last Night
Children Under 5 Slept Under Net
Number of Children Under Net Last Night
Presence of Species Malariae
Mothers Highest Educational Level
Number of Mosquito Bed Nets
Endemic Zones
Anemia Level
Region
Presence of Species Falciparum
Features
0
20
40
60
80
100
120
Relative Importance (%)

Figure 1. *ROC and AUC for the ML Models.*

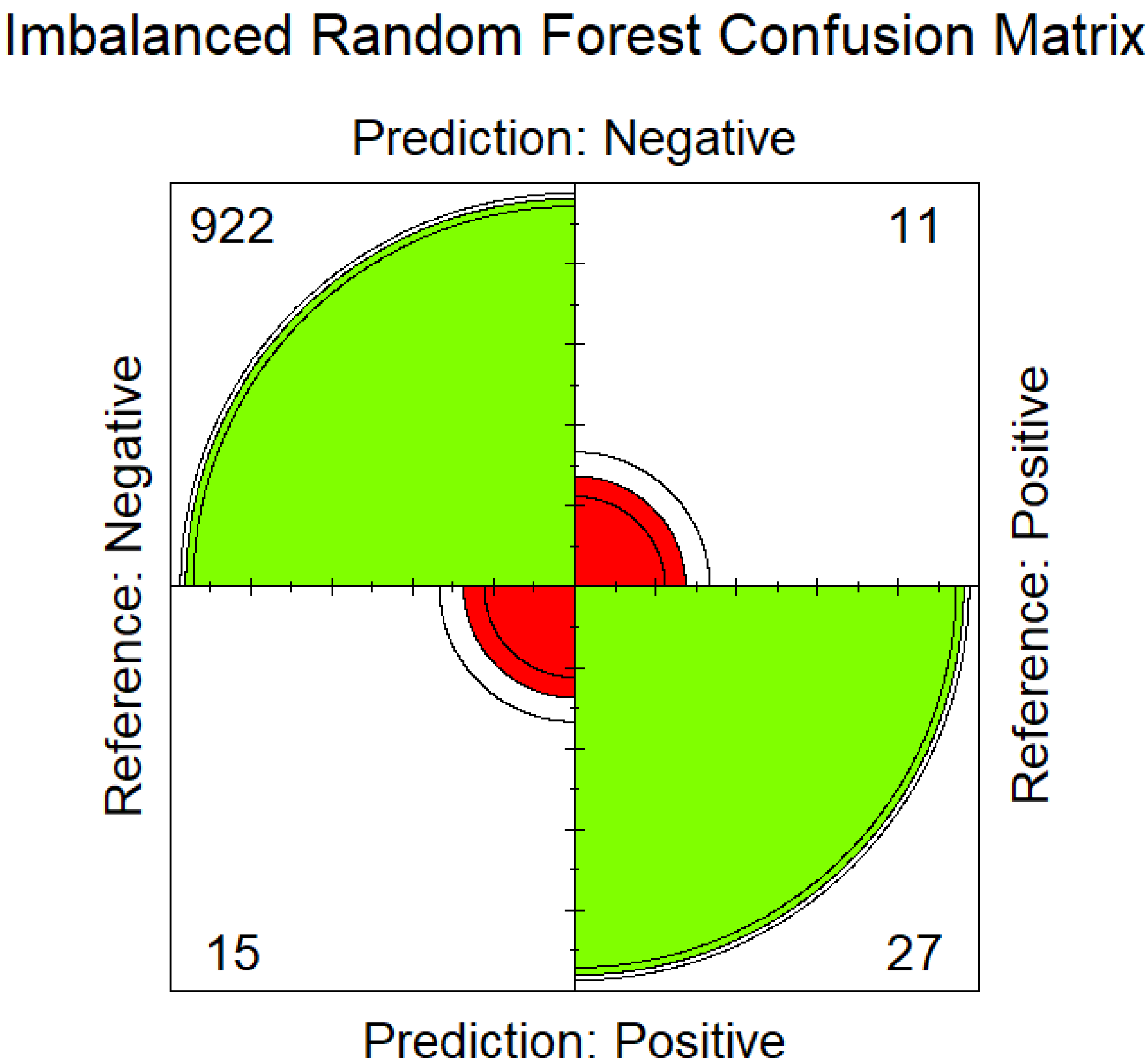
this study. Results shows that the presence of the species ***3.4. Relative Feature Importance***

Falciparum is the most important feature in the classification and prediction of malaria occurrence giving 100% relative The relative performance of feature shows the percentage importance. Having or not having mosquito net was found to contribution of feature to the variation in the outcome variable.

have 0% relative importance in classifying and predicting the Figure 2 shows the feature's relative importance from the occurrence malaria in Kenya.

random forest model which was the overall best ML model in

Figure 2. *Feature's Relative Importance Plot.*



American Journal of Theoretical and Applied Statistics

http://www.sciencepg.com/journal/ajtas ***3.5. Confusion Matrix***

resource allocation for interventions, and ultimately reducing the incidence and impact of malaria in the Kenya. From the The confusion matrix in Figure 3 shows the correctly clas- results, the target intervention, resources and funds allocation sified cases of malaria tests results and the mis-classified tests should be channeled to area with presence of plasmodium results as well. The matrix aid in the calculation of model's falciparum, regions susceptible to malaria, endemic zones and accuracy. The model's accuracy is obtained as shown in areas with higher anemic severity.

equation 18

□□□□□□□□ =

Abbreviations

□□□□ □□□□□□□□+□□□□ □□□□□□□

(18)

□□□ □□□□□□□□□+□□□□ □□□□□□□+□□□□□ □□□□□□□□+□□□□□ □□□□□□□

AUC

Area Under the Curve

ROC

Receiver Operating Characteristic

KNBS

Kenya National Bureau of Statistics

SVM

Support Vector Machine

RF

Random Forest

KNN

Kernel Nearest Neighbors

ML

Machine Learning

Acknowledgments

We would like to express our sincere gratitude to the Kenya National Bureau of Statistics (KNBS) and the Kenya National Data Archive (KeNADA) for providing the data used in this study.

Author Contributions

Dennis Muriithi: *Conceptualization, Data curation, Formal Analysis, Methodology, Writing – original draft, Writing*

– review & editing

Figure 3. *The Random Forest Confusion Matrix.*

Victor Wadera Lumumba: *Conceptualization, Data curation, Formal Analysis, Methodology, Writing – original draft, Writing – review & editing* ***4. Conclusion***

Mark Okongo: *Conceptualization, Data curation, Formal Analysis, Methodology, Writing – original draft, Writing –*

Malaria is still a killer disease globally and Kenya is not review & editing

exempted from the threat of this disease. As a result, possible measure and mitigation strategies have to be put in place to ***Funding***

address the malaria incidences in Kenya and reduce malaria related deaths among children below five years which is the This research received no external funding.

most hit category. In the five ML models estimated in this study to classify and predict the final malaria results test, random forest emerged as the most preferred model due to its ***Conflicts of Interest***

higher classification accuracy and better model performance.

The model attained a higher classification accuracy of ap-The author declares no conflicts of interest.

proximately 97.33%, with a higher sensitivity and specificity of approximately 71.1% and 98.4%, respectively. Besides, ***References***

random forest model had a relatively higher balance accuracy of approximately 84.7% and an area under the curve of 95.6%.

[1] Capili, B. (2021). Cross-Sectional Studies. The American The results indicated that the presence of plasmodium falci-Journal of Nursing/American Journal of Nursing, 121(10), 59–

parum was found to be the most important feature in classi-62. https://doi.org/10.1097/01.naj.0000794280.73744.fe fying final malaria test results, followed by region, endemic

[2] Chapelle, O. (2007). Training a Support Vector Machine in the zone, and anemic level. In conclusion, employing Machine Primal. Neural Computation, 19(5), 1155–1178.

learning algorithms enhances early detection, optimizing https://doi.org/10.1162/neco.2007.19.5.1155

[3]

Adeyemo, A. O., Aborode, A. T., Bello, M. A., Obianuju, A. F., Lazaro, S., Bailey, J. A., Juliano, J. J., Gutman, J. R., & Hasan, M. M., Kehinde, D. O., Hossain, M. S., Bardhan, M., Imisi-Ishengoma, D. S. (2023). Malaria species prevalence among oluwa, J. O., & Akintola, A. A. (2022). Malaria vaccine: The lasting asymptomatic individuals in four regions of Mainland Tanza-solution to malaria burden in Africa. Annals of Medicine and Sur-nia. MedRxiv (Cold Spring Harbor Laboratory).

gery, 79, 104031. https://doi.org/10.1016/j.amsu.2022.104031

https://doi.org/10.1101/2023.12.28.23300584

[4] Agapaki, E., & Nahangi, M. (2020). Scene understanding and

[12] Sato, S. (2021). Plasmodium—a Brief Introduction to the model generation. Elsevier EBooks, 65–167.

Parasites Causing Human Malaria and Their Basic Biology.

https://doi.org/10.1016/b978-0-12-815503-5.00003-6

Journal of Physiological Anthropology, 40(1).

https://doi.org/10.1186/s40101-020-00251-9

[5] Al-Obaidi, K. M., Ismail, M., & Malek, A. (2014). A study of the impact of environmental loads that penetrate a passive

[13] Stavropoulos, G., Voorstenbosch, R. van, Schooten, F.-J. van, skylight roofing system in Malaysian buildings. Frontiers of

& Smolinska, A. (2020). Random Forest and Ensemble Architectural Research, 3(2), 178–191.

Methods. Elsevier EBooks, 661–672.

https://doi.org/10.1016/j.foar.2014.03.004

https://doi.org/10.1016/b978-0-12-409547-2.14589-5

[6] Galal, A., Marwa Talal, & Moustafa, A. A. (2022). Applica-

[14] Takken, W. (2021). The mosquito and malaria. Routledge tions of machine learning in metabolomics: Disease modeling EBooks, 109–122. https://doi.org/10.4324/9781003056034-11

and classification. Frontiers in Genetics, 13.

https://doi.org/10.3389/fgene.2022.1017340

[15] Trampuz, A., Jereb, M., Muzlovic, I., & Prabhu, R. M. (2003).

Clinical review: Severe Malaria. Critical Care, 7(4), 315.

[7] Giesen, C., Roche, J., Redondo-Bravo, L., Ruiz-Huerta, C., https://doi.org/10.1186/cc2183

Gomez-Barroso, D., Benito, A., & Herrador, Z. (2020). The impact of climate change on mosquito-borne diseases in Africa.

[16] WHO. (2024). Malaria. WHO | Regional Office for Africa.

Pathogens and Global Health, 114(6), 1–15.

https://www.afro.who.int/health-topics/malaria https://doi.org/10.1080/20477724.2020.1783865

[17] Cunningham, P., & Delany, S. J. (2007, April 27). k-Nearest

[8] Ileperuma, K., Jampani, M., Sellahewa, U., Panjwani, S., & neighbor classifiers. ResearchGate; Association for Compu-Amarnath, G. (2023). Predicting Malaria Prevalence with ting Machinery.

Machine Learning Models Using December 2023 Colombo, https://www.researchgate.net/publication/228686398_k-Neare Sri Lanka. https://www.iwmi.cgiar.org/Publications st_neighbour_classifiers

[9] Lee, Y. W., Choi, J. W., & Shin, E.-H. (2021). The machine

[18] Kazeem, I., & Adebanji, S. (2021, November 22). A model for learning model for predicting malaria using clinical infor-predicting malaria outbreak using machine learning technique.

mation. Computers in Biology and Medicine, 129, 104151.

ResearchGate; Scientific Annals of Computer Science.

https://doi.org/10.1016/j.compbiomed.2020.104151

https://www.researchgate.net/publication/356439342

[10] Oladipo, H. J., Tajudeen, Y. A., Oladunjoye, I. O., Yusuff, S. I.,

[19] World. (2023, December 4). Malaria. Who.int; World Health Yusuf, R. O., Oluwaseyi, E. M., AbdulBasit, M. O., Adebisi, Y.

Organization: WHO.

A., & El-Sherbini, M. S. (2022). Increasing challenges of ma-https://www.who.int/news-room/fact-sheets/detail/malaria laria control in sub-Saharan Africa: Priorities for public health

[20] Owoko, L. (2024, June 11). Kenya's child malaria deaths fall research and policymakers. Annals of Medicine and Surgery, three-fold on campaigns. Business Daily; Business Daily.

81(104366). https://doi.org/10.1016/j.amsu.2022.104366

https://www.businessdailyafrica.com/bd/corporate/health/ken

[11] Popkin, Z. R., Seth, M. D., Madebe, R. A., Rule Budodo, ya-s-child-malaria-deaths-fall-three-fold-on-campaigns—4654

Bakari, C., Francis, F., Dativa Pereus, Giesbrecht, D. J., 574

Mandara, C. I., Mbwambo, D., Aaron, S., Abdallah Lusasi,

[21]