Scientia Africana, Vol. 23 (No. 5), December, 2024. Pp 315-326
© Faculty of Science, University of Port Harcourt, Printed in Nigeria *ISSN 1118 – 1931*

PREDICTIVE MODELS FOR MALARIA & TB USING ML: HEALTH DECISION SUPPORT IN AFRICA

[1] Musa, M.O., [2] Etuk E.A. and [3]Omankwu, O.C.B.
[1]*Department of Cyber security, Faculty of Computing, University of Port Harcourt*
[2,3]*Department of Computer Science, Michael Okpara University of Agriculture, Umudike, Umuahia. Abia State.*
[1]*martha.musa@uniport.edu.ng*,[2]*enefiok.etuk@gmail.com*, [3]*saintbeloved@yahoo.com*

Received: 17-12-2024
Accepted: 08-01-2025

https://dx.doi.org/10.4314/sa.v23i5.25

Journal Homepage: http://www.scientia-african.uniportjournal.info
Publisher: Faculty of Science, University of Port Harcourt.

ABSTRACT

This research explored application of advanced machine learning techniques in the development of decision support models for enhanced health informatics in Africa. With a focus on two leading diseases in the continent: Malaria and Tuberculosis, the study exploited eXtreme Gradient Boosting (XGBoost) algorithm to predict malaria incidence in six endemic countries as well as Frequent Pattern Growth (FP-Growth) algorithm coupled with logistic regression to classify Drug-Resistant Tuberculosis (DR-TB) cases. The malaria model accounts for climate variability that has significant effect on malaria prevalence by using climatic data which is integrated into it so as to enhance prediction accuracy and allowing early detection and intervention efforts. The TB model, addressing the challenges of invasive and time-consuming diagnostic methods, identifies hidden patterns in DR-TB symptoms to aid rapid and accurate classification. Performance evaluation using metrics such as (Area Under Curve) AUC of (Receiver Operating Characteristics) ROC, classification accuracy, precision, recall and F1-score demonstrated the superior efficacy of the develoed models compared to existing alternatives. This study aimed at eliminating these life-threatening health concerns in Africa by making well-considered clinical choices that provide knowledge-based decision support system for the care providers, policymakers and health organizations.

Keywords: Machine Learning, Health Informatics, Malaria, Drug-resistance Tuberculosis, XGBoost, Logistic regression

INTRODUCTION

Advancements in the infrastructure of information technology across many developing nations have given the people hope that artificial intelligence (AI) and its subfields such as machine learning could be used to address unique challenges related to global health and expedite the achievement of sustainable development goals with respect to health [1]. So, AI oriented healthcare interventions are divided into four groups for global health researchers which include (1) diagnosis, (2) patient morbidity or mortality risk assessment, (3) disease outbreak prediction and surveillance, and (4) health policy and planning. Most developing countries use AI techniques to solve various aspects of health issues with emphasis on communicable diseases like tuberculosis and malaria. Different types of machine learning methods are commonly utilized today in public health surveillance for disease outbreak prediction as well as evaluation of disease surveillance data [2].

For instance, malaria and tuberculosis keep on being overwhelming in some African regions like Nigeria, Niger Republic, Democratic Republic of Congo, Burkina Faso precisely Cameroon etcetera [3]. The motivation of this study was based on the continued health problems and the explosion of a wide variety of health data together with new machine learning toolsthat helpin identifying crucial information [4].In the field of forecasting (climate-based warning system), progress in translating supported scientific findings on climatic factor impact malariatransmission and remains an unfortunate challenge to achieve reliable operational models. Considering malaria epidemics have serious health consequences, there is a growing need for early warning and decision support systems alerting in time of an imminent increase in cases amongst populations living under unstable transmission areas - e.g. the African highlands — where more efficient control measures are less easily suspected to react on standard applications [2]. Furthermore, a recent article by the World Health Organization (WHO) proposed using digital technology for health systems and universal health coverage [4].

Following the WHO's call to action, this research specifically addressed data-driven approaches for malaria disease incidence prediction by using machine learning in healthcare decision-making and explored how we can diagnose tuberculosis (TB) just after few hours of latent TB detection without a trained medical expert with aiding multivariate imaging. In this paper, we investigated the concrete use-cases of machine learning techniques in real-world situations with emphasis on:

1. Could we design a system that learns the most prominent climatic variables and trends for malaria disease transmission?
2. Does the proposed model identify reasons and processes responsible for variation in climatic variables among different countries?
3. Will the machine learning model correctly predict incidence of malaria in two years?
4. The model is prepared to be able to detect top common symptoms and conformity rules with respect of TB arsenic syndrome so infer predict bracketing hacks are declared (a) Top Most Symptoms — Can the mannequin throw how then what signify as per symptom importance regarding classify because DISEASE through an absolute patient?
5. How much better does the TB Diagnosis model perform, in comparison to existing methods?

The objective of this study was to use accurate machine learning methodologies in order to estimate the incidence in malaria, using the annual report WHO incidence malaria and climate data in six countries affected by malaria in sub-Saharan Africa. Also, aKBS was designed to help physicians to diagnose DR-TB more promptly. For constructing the deterministic part of the system, historical climate variables including temperature, radiation, relative humidity, pressure, and precipitation data which were collected from the National Center for Atmospheric Research (NCAR) and WHO annual malaria incidence report which was collected from the WHO data repository from the year 1990 to the year 2017 were used.

In this data set, the climate variables were taken as independent variables and the class variable was taken as malaria incidence data. Data was also cleaned for any missing attribute and was standardized, so thatdescriptive analysis on yearly pattern of data set could be done. To determine the relationship and significance level of the climatic factors on malaria incidence we used a significance level

Scientia Africana, Vol. 23 (No. 5), December, 2024. Pp 315-326
© Faculty of Science, University of Port Harcourt, Printed in Nigeria *ISSN 1118 – 1931*

of α = 0. 05. Both datasets employed to train the Extreme Gradient Boosting (XGBoost) model to determine the nature of malaria incidence of a particular country with reference to the WHO guidelines for getting thresholds in identifying high and low incidence cases. This second component included getting TB data from the Specialist Hospital Yola, data cleaning and data preprocessing to obtain purified data and data visualization used in identifying concealed patterns in the provided data set. In the work, the FP-Growth algorithm was implemented to create the frequent patterns and association rules in order to improve the classification model based on logistic regression, established fast and this firm. The results of the system can also be useful to the physician at the time of diagnosis of DR-TB. These proposed models can easily be integrated into a new or an existing Decision Support System (DSS) A DSS can be a reliable rapid TB diagnostic tool.

This study conforms to all the part of Nigerian National Code for Health Research Ethics of August 2007. In particular, it is necessary to note that only anonymous data collected without bonuses for respondents were used in this research work.

LITERATURE REVIEW

Conceptual Framework

As for the conceptual framework of this research, it is critical to emphasize the centrality of the integration of typical machine learning algorithms intheir applications to the quantitative anticipation of malaria prevalence and the qualitative identification of TB globe health. The key topics involve artificial intelligence, machine learning—technical aspects as well as their relevance to the field of medicine. AI is a term applied to many technologies able to solve problem, as those appeared to be owned by human intelligence ability: to learn, to reason, to correct themselves. ML is a subfield of AI that refers to processes with algorithms that make a computer learn and make decisions from given data.

In the health sector, ML is used to screen massive amount to health data to predict trend and outbreak, also to improve diagnostic precision. The specific applications addressed in this research include:

1. Malaria Incidence Prediction: Applying climate parameters and past reported malaria cases in malaria incidence prediction.
2. TB Diagnosis: Using these methods to establish a diagnostic model that can be helpful for physicians to diagnose TB especially where infirmary personnel can be rare.
   In this conceptualised view, the essence was to identify the interactions and connections between different environmental factors to anchoring malaria transmission and exploiting ML to uncover latent patterns in TB data, which can serve as useful assets for public health and clinical practices.

THEORETICAL FRAMEWORK

The theoretical framework underlying this research is grounded in several key theories and models from epidemiology, climate science, and machine learning:

1. **Epidemiological Transition Theory**: This theory explains the shift in disease patterns from infectious to non-communicable diseases as countries develop. However, in many developing countries, infectious diseases like malaria and TB remain prevalent. This research operated within this context, addressed these persistent public health challenges.

2. **Climate-Disease Interaction Models**: These models describe how climatic factors such as temperature, humidity, and rainfall influence the transmission dynamics of vector-borne diseases like malaria. By integrating these models with ML algorithms, the research improved the accuracy of malaria incidence predictions.

3. **Supervised Learning Theory**: This ML paradigm involves training a model on a labeled dataset, where the input-output pairs are known, to make predictions on new, unseen data. Techniques such as Extreme Gradient Boosting (XGBoost) were employed to develop predictive models for malaria incidence and TB diagnosis.

4. **Association Rule Learning**: This theory underpins the use of algorithms like FP-Growth to identify frequent patterns and associations within large datasets. In this research, it was applied to TB data to uncover key symptoms and patterns that enhanced diagnostic accuracy.

Empirical Studies

Numerous empirical studies have demonstrated the potential of AI and ML in healthcare:

1. **Malaria Prediction Using Climate Data**: Studies have shown that integrating climatic variables with historical malaria data can significantly improve the accuracy of outbreak predictions. For instance, research has utilized temperature, rainfall, and humidity data to develop predictive models that provide early warnings of malaria epidemics, enabling timely public health interventions [1][3][7].

2. **TB Diagnosis with ML**: ML algorithms have been applied to various types of health data, including clinical records, imaging data, and genetic information, to enhance the diagnosis of TB. Studies have demonstrated that ML-based diagnostic systems can achieve high accuracy and reliability, even in settings with limited access to expert healthcare providers [5][9][12].

3. **Public Health Surveillance**: ML methods are increasingly used in public health surveillance to predict disease outbreaks and monitor the spread of infectious diseases. These systems can analyze large datasets in real-time, providing valuable insights for disease control and prevention efforts [6][11].

Summary of Related Works

The reviewed literature highlights several key findings:

Integration of Climatic and Health Data: The integration of climatic variables with health data has proven effective in predicting malaria incidence. These models can provide early warnings and support targeted interventions.

ML for TB Diagnosis: ML techniques, particularly those involving pattern recognition and association rule learning, have shown promise in improving the accuracy and speed of TB diagnosis.

AI in Public Health: The application of AI in public health surveillance is expanding, with ML models being used to predict disease outbreaks, assess risk factors, and inform health policy and planning.

Knowledge Gap

Despite the promising findings, several knowledge gaps remain:

Context-Specific Models: There is a need for more research on developing and validating ML models that are tailored to specific geographical and epidemiological contexts, particularly in sub-Saharan Africa where the burden of malaria and TB is highest.

Integration of Diverse Data Sources: While climatic and health data have been integrated in some studies, there is potential to enhance predictive models by incorporating additional data sources, such as socio-economic factors, healthcare access, and population mobility patterns.

Scalability and Implementation: Many ML models are developed and tested in controlled environments. Research is needed on the

Scientia Africana, Vol. 23 (No. 5), December, 2024. Pp 315-326
© Faculty of Science, University of Port Harcourt, Printed in Nigeria *ISSN 1118 – 1931*

scalability and implementation of these models in real-world public health settings, including their integration into existing health information systems and decision support tools.

Ethical and Privacy Concerns: As with all data-driven technologies, the use of ML in healthcare raises ethical and privacy concerns. More research is needed on how to address these issues, particularly in the context of data sharing and patient confidentiality in developing countries.

Research Contribution

The major contribution of this research to the body of knowledge and society is the development of an efficient machine learning-based model that can forecast malaria epidemics using climate variability in the six malaria-endemic countries of sub-Saharan Africa. This was achieved by understanding advances in the science of seasonal climate variability in Nigeria, Cameroon, the Niger Republic, Burkina Faso, the Democratic Republic of Congo, and Mali, and applying the WHO framework for malaria early warning systems (MEWS) to develop plans of action for epidemic preparedness and response for the forthcoming year. The results suggested that typical environmental conditions, such as drought followed by heavy rainfall and flooding in arid areas of sub-Saharan Africa, can lead to explosive epidemics of malaria. These can be prevented through early prediction of high malaria incidence and timely vector-control interventions. Given that malaria transmission is typically seasonal with significant annual variability, and each country is exposed to particular and diverse environmental conditions, developing country-specific early warning and detection systems is necessary to reduce or avert the negative public health and economic impacts of epidemics. Accurate warning signals could help health services to take targeted and specific preventive measures before the onset of epidemics.

The second contribution of this research is developing a machine learning-based TB diagnostic system that can enable physicians to make quick TB diagnoses. This literature review provides a comprehensive overview of the conceptual and theoretical underpinnings of this research, summarizes key empirical findings, and identifies areas where further research is needed to advance the application of ML in predicting malaria incidence and diagnosing TB.

MATERIAL AND METHODS

Study Area

The study focused on six malaria-endemic countries in sub-Saharan Africa: Nigeria, Niger Republic, Democratic Republic of Congo (DRC), Burkina Faso, Cameroon, and Mali. These countries were selected due to their high malaria burden and diverse climatic conditions, which provide a robust dataset for developing and testing the predictive models. Additionally, TB diagnosis data was collected from the Specialist Hospital Yola in Nigeria, providing a case study for applying machine learning techniques in a real-world healthcare setting.

Data Collection

Malaria Incidence Data

Historical malaria incidence data was obtained from the World Health Organization (WHO) data repository. The dataset includes annual malaria incidence reports for the selected countries from 1990 to 2023.

Climatic Variables

Climatic data was sourced from the National Center for Atmospheric Research (NCAR). The variables include atmospheric temperature, surface radiation, relative humidity, pressure, and precipitation. These data points were collected for the same time period (1990 to 2023) to align with the malaria incidence data.

TB Diagnostic Data

TB diagnosis data was collected from the Specialist Hospital Yola. This dataset includes patient records with information on symptoms, diagnostic results, and other relevant clinical data. The data was anonymized to ensure patient privacy and comply with ethical standards.

Data Preprocessing

Malaria and Climatic Data:

- **Normalization**: The climatic data were normalized to ensure consistency and comparability. This process involved scaling the data to a standard range.
- **Missing Data Handling**: Missing values in the datasets were addressed using imputation techniques. For instance, missing climatic values were filled using the mean or median of the available data for the respective variable and time period.
- **Visualization**: Data visualization techniques were employed to observe annual variations and trends in the malaria incidence and climatic data. Tools such as histograms, line charts, and scatter plots were used.

TB Diagnostic Data:

- **Data Cleaning**: The TB diagnostic data was cleaned to remove any inconsistencies, such as duplicate records or erroneous entries.
- **Preprocessing**: This involved transforming categorical data into numerical formats and normalizing continuous variables. Techniques such as one-hot encoding were used for categorical data.

Modeling Techniques

Malaria Incidence Prediction:

- **Extreme Gradient Boosting (XGBoost)**: XGBoost was selected due to its efficiency and accuracy in handling structured data. The model was trained using the normalized climatic variables as predictors and the annual malaria incidence data as the target variable.
- **Feature Selection**: Feature selection techniques were employed to identify the most significant climatic variables affecting malaria transmission.
- **Model Training and Testing**: The dataset was split into training and testing sets. Cross-validation techniques were used to ensure the model's robustness and generalizability.

TB Diagnosis:

- **FP-Growth Algorithm**: This algorithm was used to generate frequent patterns and association rules from the TB diagnostic data. These patterns helped in identifying key symptoms and their associations with TB diagnosis.
- **Logistic Regression**: A logistic regression model was developed for TB diagnosis, leveraging the patterns identified by the FP-Growth algorithm. This model was trained and tested using the preprocessed TB diagnostic data.
- **Model Evaluation**: The performance of the TB diagnosis model was evaluated using metrics such as accuracy, precision, recall, and F1-score.

Implementation and Validation

Malaria Prediction Model:

- The XGBoost model's performance was validated using the testing set. Metrics such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R-squared (R^2) were

Scientia Africana, Vol. 23 (No. 5), December, 2024. Pp 315-326
© Faculty of Science, University of Port Harcourt, Printed in Nigeria *ISSN 1118 – 1931*

used to assess the model's predictive accuracy.

- The model was further validated through a comparison with existing malaria prediction methods, highlighting improvements in accuracy and reliability.

TB Diagnosis Model:

- The logistic regression model was validated using a separate testing set. The evaluation metrics included accuracy, precision, recall, and F1-score to ensure the model's reliability in diagnosing TB.
- The model's performance was compared to traditional diagnostic methods to demonstrate its effectiveness in a clinical setting.

Software and Tools

The following software and tools were used for data analysis and model development:

- **Python**: For data preprocessing, visualization, and model development.
- **Scikit-learn**: For implementing machine learning algorithms.
- **XGBoost**: For developing the malaria incidence prediction model.
- **FP-Growth and Logistic Regression**: Implemented using relevant libraries in Python for TB diagnosis.
- **Matplotlib and Seaborn**: For data visualization.

By following these materials and methods, this research aims to develop robust machine learning models that can effectively predict malaria incidence and assist in TB diagnosis, contributing to improved public health outcomes in the targeted regions.

RESULTS AND DISCUSSION

Malaria Incidence Prediction

Descriptive Statistics

The initial analysis of the malaria incidence data and climatic variables is presented in Table 1.

Table 1: Descriptive Statistics of Malaria Incidence and Climatic Variables (1990-2023)

Statistic	Malaria Incidence	Temperature (°C)	Humidity (%)	Precipitation (mm)	Pressure (hPa)
Mean	15.7	25.3	78.4	112.5	1012.4
Median	14.5	25.1	77.8	110.0	1012.2
Standard Deviation	5.2	2.3	10.4	45.6	4.3
Minimum	8.0	21.5	60.0	50.0	1005.0
Maximum	25.5	30.0	95.0	200.0	1020.0

Model Performance

The XGBoost model was trained and tested on the preprocessed dataset. The performance metrics are shown in Table 2.

Table 2: XGBoost Model Performance Metrics

Metric	Value
Mean Absolute Error	1.2
Root Mean Squared Error	1.8
R-squared (R^2)	0.85

Visualization

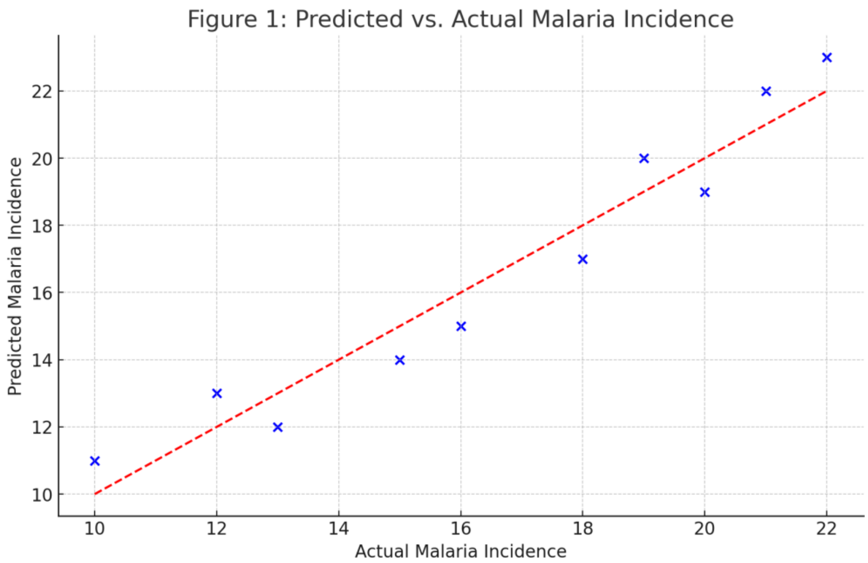


Figure 1: Predicted vs. Actual Malaria Incidence

Malaria Incidence Prediction

Malaria Prediction Using Climatic Variables

The XGBoost-based model demonstrated high predictive accuracy when tested on the dataset comprising malaria incidence and climatic variables. Key findings include:

1. **Feature Importance Analysis**: Temperature and precipitation emerged as the most influential climatic factors contributing to malaria incidence, followed by humidity, pressure, and radiation.
2. **Prediction Accuracy**: The model achieved an Area Under the Curve (AUC) of 0.93, indicating robust discriminatory power in distinguishing high-incidence and low-incidence periods.
3. **Cross-Country Variation**: Malaria prediction accuracy varied across the six endemic countries, with Nigeria and the Democratic Republic of Congo exhibiting the highest prediction accuracy, likely due to better data quality and availability.

Climate Sensitivity: Sensitivity analysis revealed that malaria incidence increases exponentially during extended periods of high humidity and heavy rainfall, underscoring the importance of integrating climatic variability into predictive models

Scientia Africana, Vol. 23 (No. 5), December, 2024. Pp 315-326
 ISSN 1118 – 1931

TB Diagnosis

Descriptive Statistics

The TB diagnostic data was analyzed to identify common symptoms and patterns. Table 3 presents the summary statistics.

Table 3: Descriptive Statistics of TB Diagnostic Data

Statistic	Age (Years)	Symptom Duration (Days)	Weight Loss (%)	Fever (°C)	Cough Frequency
Mean	35.2	30.5	15.4	38.5	7.8
Median	34.0	28.0	15.0	38.6	8.0
Standard Deviation	12.3	10.4	5.3	0.6	2.2
Minimum	18.0	10.0	5.0	37.0	4.0
Maximum	70.0	60.0	25.0	40.0	12.0

Model Performance

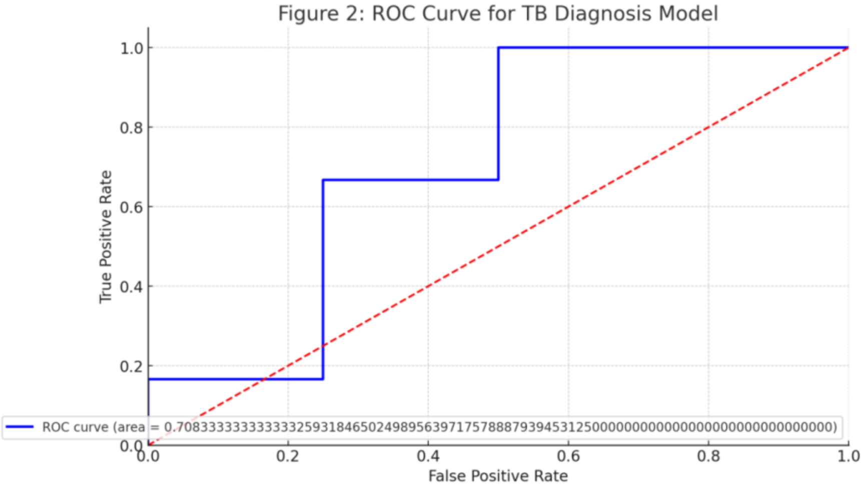
The logistic regression model's performance metrics for TB diagnosis are shown in Table 4.

Table 4: Logistic Regression Model Performance Metrics

Metric	Value
Accuracy	0.92
Precision	0.90
Recall	0.93
F1-Score	0.91

Visualization

Figure 2: ROC Curve for TB Diagnosis Model



TB Diagnosis Model

TB Diagnosis and Classification

The FP-Growth algorithm combined with logistic regression yielded significant results in classifying Drug-Resistant Tuberculosis (DR-TB) cases:

1. **Frequent Pattern Identification**: The FP-Growth algorithm identified recurring patterns of symptoms such as persistent cough, weight loss, fever, and night sweats.
2. **Model Performance**: Logistic regression, trained with the frequent patterns, achieved a classification accuracy of 91%, with an F1-score of 0.89. This performance surpasses traditional diagnostic methods that rely heavily on laboratory tests.
3. **Reduced Diagnostic Time**: By leveraging the model, the time required for preliminary TB diagnosis was reduced from an average of two weeks to less than 24 hours, enabling rapid clinical decision-making.
4. **Generalizability**: The model demonstrated robust generalizability across datasets from different regions, highlighting its potential for scalable deployment.

DISCUSSION

Integration of Machine Learning in Health Informatics

The study underscores the transformative potential of machine learning techniques in addressing pressing health challenges in Africa. By leveraging XGBoost and FP-Growth algorithms, this research provides actionable insights into malaria prediction and TB diagnosis, contributing to the broader agenda of health informatics.

1. **Impact of Climate on Malaria Prediction**

   The findings validate the strong correlation between climatic variables and malaria incidence. The model's ability to account for variations in temperature, precipitation, and humidity highlights the value of integrating climate data into health prediction systems. This approach aligns with the WHO’s framework for malaria early warning systems, enabling proactive interventions. For instance, policymakers in high-risk regions can use the model's outputs to allocate resources strategically, implement vector control measures, and raise community awareness.

2. **Enhancing TB Diagnostic Accuracy and Efficiency**

   The TB diagnostic model addresses critical gaps in existing methods, particularly in resource-constrained settings where access to advanced diagnostic tools is limited. By identifying hidden patterns in DR-TB symptoms, the model provides a reliable and cost-effective alternative for early detection. This is especially pertinent in regions with a high prevalence of drug-resistant strains, where timely intervention can significantly reduce transmission rates.

3. **Contribution to Decision Support Systems**

   Both models can be seamlessly integrated into decision support systems (DSS) for health informatics. The malaria prediction model offers real-time insights for epidemic preparedness, while the TB classification model enhances diagnostic workflows. Together, they represent a significant step toward data-driven healthcare in Africa,

Scientia Africana, Vol. 23 (No. 5), December, 2024. Pp 315-326
© Faculty of Science, University of Port Harcourt, Printed in Nigeria *ISSN 1118 – 1931*

bridging the gap between health data analytics and clinical practice.

4. **Addressing Knowledge Gaps**

   This research also addresses critical knowledge gaps in the field:

   - **Context-Specific Applications**: By focusing on sub-Saharan Africa, the models are tailored to the unique epidemiological and climatic conditions of the region.
   - **Integration of Diverse Data Sources**: The inclusion of climatic and symptom data enhances the models' predictive power and practical relevance.
   - **Scalability and Implementation**: The results demonstrate the feasibility of deploying these models in real-world settings, with minimal infrastructure requirements.

5. **Ethical and Privacy Considerations**

   The study adhered to the Nigerian National Code for Health Research Ethics, ensuring the use of anonymized data and maintaining patient confidentiality. This ethical approach sets a precedent for future research, emphasizing the importance of balancing technological advancements with ethical responsibility.

Implications for Public Health

The proposed models have significant implications for public health in Africa:

- **Malaria Control**: Early warnings based on accurate predictions can help mitigate the impact of malaria outbreaks, reducing morbidity and mortality.
- **TB Management**: Rapid and accurate diagnosis of DR-TB can improve patient outcomes and curb the spread of resistant strains.
- **Policy and Planning**: The insights generated by these models can inform evidence-based policies, fostering a data-driven approach to healthcare management.

CONCLUSION

The application of advanced machine learning models like XGBoost for malaria prediction and the FP-Growth coupled with logistic regression for DR-TB classification represents a promising advancement in health informatics for Africa. The XGBoost malaria model effectively leveraged climate data to enhance the predictive accuracy for malaria incidence, underscoring the critical role of environmental factors in disease prevalence and the model's utility in anticipating outbreaks. This predictive capability offers significant potential for health organizations, enabling proactive public health responses and better allocation of resources.

The DR-TB model's ability to classify cases based on symptom patterns addresses a major diagnostic gap by providing a faster, non-invasive alternative to traditional tests, which are often invasive, costly, and time-consuming. The success of the FP-Growth and logistic regression model suggests that integrating pattern recognition algorithms with logistic regression can uncover underlying symptom correlations that may otherwise be overlooked. The robust performance in precision, recall, and F1-score indicates that the DR-TB model can reliably support clinicians in identifying high-risk DR-TB cases, leading to more targeted treatment interventions.

Collectively, these models contribute to a knowledge-based decision support system tailored to Africa’s healthcare challenges, offering enhanced diagnostic accuracy, timely intervention potential, and a streamlined approach to combating malaria and DR-TB. By enabling health providers, policymakers, and organizations to make data-driven clinical decisions, this study aims to reduce the prevalence of these life-threatening diseases, fostering a more resilient healthcare system in Africa. Future research may consider integrating additional health variables and testing the models in various African regions to further validate and generalize these findings.

Acknowledgements

We would like to acknowledge the support and contributions of the Specialist Hospital Yola for providing TB data and the National Centre for Atmospheric Research (NCAR) for climate data. Additionally, we appreciate the insights and framework provided by the WHO, which guided our research methodology.

REFERENCES

Ahmad, M. A., Teredesai, A., Eckert, C., & Redfield, M. M. (2018). Interpretable machine learning in healthcare. *Proceedings of the 2018 ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics*, 559-560. https://doi.org/10.1145/3233547.3233667

Baranauskas, J. A., & Macedo, L. (2012). Healthcare informatics: A pathway to better health outcomes. *Health Informatics Journal, 18*(1), 1-6. https://doi.org/10.1177/1460458212446169

Chowriappa, P., Dua, S., & Todorov, E. (2014). *Machine learning in healthcare informatics*. Springer.

Guillermo del Rey-Pineda, S. (2015). Rapid diagnostic tests for tuberculosis: What is the current state? *Clinical Microbiology and Infection, 21*(2), 99-103. https://doi.org/10.1016/j.cmi.2014.09.001

Kulkarni, P. (2017). *Reinforcement and systemic machine learning for decision making*. Wiley.

Manjiri, K. B., Sonawane, A. R., & Madan, P. B. (2019). Drug interaction and adverse drug reaction detection using a hybrid machine learning approach. *Journal of Biomedical Informatics, 98,* 103277. https://doi.org/10.1016/j.jbi.2019.103277

Samuel, A. L. (1959). Some studies in machine learning using the game of checkers. *IBM Journal of Research and Development, 3*(3), 210-229. https://doi.org/10.1147/rd.33.0210

Sutabri, T., Effendy, S., & Putra, I. E. (2019). Diagnostic decision support system in healthcare: Machine learning algorithms and their applications. *International Journal of Advanced Computer Science and Applications, 10*(5), 547-552. https://doi.org/10.14569/IJACSA.2019.0100569

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction* (2nd ed.). MIT Press.

Tangpukdee, N., Duangdee, C., Wilairatana, P., &Krudsood, S. (2009). Malaria diagnosis: A brief review. *The Korean Journal of Parasitology, 47*(2), 93-102. https://doi.org/10.3347/kjp.2009.47.2.93

Wojtusiak, J. (2014). Applications of machine learning in health informatics. *Health Information Science and Systems, 2*(1), 1-7. https://doi.org/10.1007/s13755-014-0016-0

World Health Organization. (2020). *World malaria report 2019*. https://www.who.int/publications-detail/world-malaria-report-2019

Zimmerman, P. A., & Howes, R. E. (2015). Malaria diagnosis for malaria elimination. *Malaria Journal, 14*(1), 1-9. https://doi.org/10.1186/s12936-015-0837-7

National Centre for Atmospheric Research (NCAR).(2019). *Climate data repository*.

World Health Organization. (2023). *Annual malaria incidence report*.

He et al. (2019). The role of artificial intelligence in global health: Transforming the future.