Bol. Soc. Paran. Mat.
©SPM – E-ISSN-2175-1188
SPM: www.spm.uem.br/bspm

(3s.) **v. 2026 (44) 8** : 1–23.
ISSN-0037-8712
doi:10.5269/bspm.79633

AI-Powered Mortality Prediction for HIV/AIDS Patients on ART

Okechukwu J. Obulezi*, Mohamed A. F. Elbarkawy, Mmesoma P. Nwankwo, EbereChukwu Q. Chinedu, Chinyere P. Igbokwe, Heba E. Salem and Mustafa Bayram

ABSTRACT: This paper evaluates various parameter estimation methods for the Wald distribution using simulations and real-world datasets. Simulation results confirm that increased sample sizes improve estimates, reducing bias and Root Mean-Squared Error (RMSE). The Maximum Likelihood Estimator (MLE) is generally the most robust method for large samples but unstable and biased for smaller ones, particularly in estimating λ. The Maximum product of spacing estimation (MPS) method performs well asymptotically, with bias and RMSE decreasing as sample size increases. Least Squares (LS) and Weighted Least Squares (WLS) are suitable alternatives to MLE for small-to-moderate samples, showing similar estimates and reduced bias with larger samples. The Cramer-von Mises Estimator (CvM) displayed the worst efficiency due to high RMSE. Bayesian Estimators (BE) showed greater bias and lower efficiency than their frequentist counterparts, especially for λ, with performance strongly dependent on prior selection. Application to real datasets (HIV/AIDS mortality, COVID-19 death rates, and industrial gauge measurements) demonstrates the Wald distribution's feasibility in diverse health data analysis and reliability studies. The study concludes that MLE and MPSE are efficient estimation methods, suggesting the Wald distribution is a strong candidate for applied parameter estimation. Future work could focus on improving Bayesian methods via better prior selection.

Keywords: HIV/AIDS prediction, Machine Learning, Deep Learning, mortality rates, Nigeria healthcare.

Contents

1. Introduction

Human Immunodeficiency Virus (HIV) remains a significant global health challenge, continually weakening the immune system and leading to Acquired Immunodeficiency Syndrome (AIDS) if left untreated, leaving individuals highly susceptible to opportunistic infections and opportunistic cancers. The World Health Organisation (WHO) in 1981 classified HIV/AIDS as a global health emergency, and the pandemic has unfortunately claimed over 40.4 million lives worldwide. As of 2022, an estimated 39 million people lived with HIV globally, with two-thirds of these residing in the WHO African region [2]. Ambitious 2025 global goals of 95% of all people living with HIV (PLHIV) being diagnosed, 95% of diagnosed individuals receiving life-saving Antiretroviral Therapy (ART), and 95% of those on treatment having viral suppression [2,15]. Despite significant advances in ART, which have transformed HIV to a chronic disease, the epidemic remains a serious challenge, particularly in those high prevalence and under-resourced countries with limited healthcare facilities, which hampers the speed to end AIDS by 2030 due to ongoing new infections among the youth [18]. Nigeria, for instance, records 190,950 new HIV/AIDS infections annually, making it the second highest globally [3], with an estimated 1,910,405 people living with HIV in 2023 [15]. The Nigerian HIV epidemic is generalized and affects all population groups but with profound regional variations in prevalence [9,31], and there are specific challenges like increasing mother-to-child transmission [1], calling for enhanced monitoring of disease progression and response to treatment [30]

* Corresponding author.
2020 *Mathematics Subject Classification:* 62F10.
Submitted October 18, 2025. Published March 12, 2026

and particularly in view of the large economic benefits of investing to end AIDS [19]. Incidence, deaths, and disability-adjusted life years (DALY) trends of HIV in regions like the Middle East and North Africa are ominously on the rise, indicative of the imperative need for more prevention and treatment care globally [21].

The push for this research stems directly from the basic necessity to enhance mortality prediction in HIV/AIDS patients receiving ART in Nigeria since early and proper identification of high-risk patients is key to ensuring optimal patient management, enabling focused clinical interventions, and securing proper utilization of healthcare resources. Machine learning (ML) and deep learning (DL), the powerful domains of artificial intelligence (AI), offer cutting-edge computational methods best suited to address this complex issue, best suited to identify subtle patterns and make robust predictions from large and complex datasets without explicit coding [4,22]. The primary contribution of AI in public health has been recognized widely, most notably in predicting the outbreak of the new coronavirus [5], with AI techniques, including ML and DL, used to detect emerging outbreaks [8] and utilized as major components in effective pandemic mitigation, highlighting the adoption of advanced digital tools for surveillance, forecasting, and diagnosis [7] and poised to play a crucial role in policymaking and resource allocation for disproportionately affected communities [6]. Past studies have explored ML for forecasting HIV transmission among high-risk sub-Saharan African areas [10,11,12,20] with consistent evidence of the potential of ML to identify high-risk areas and forecast outbreak [14]. Overall, ML approaches have been shown greater efficacy in identifying individuals at risk of HIV acquisition than traditional approaches in generalized epidemic settings [14]. Bayesian predictive modeling has also been applied to estimate HIV burden and prevalence in Nigeria [16], while mixed-effect models have been utilized to estimate HIV testing and condom use coverage toward UNAIDS targets [17,13]. Moreover, artificial neural networks (ANNs) have also demonstrated enormous potential in predictive modeling and medical diagnosis in HIV research [23,26], with high precision in predicting adolescent HIV infection in Nigeria [29] and being effective for viral load and CD4 staging among adults undergoing ART based on various ML algorithms like XGBoost and Gradient Boosting [25,30,27].

The novelty of the study is in its integrative and contextualized approach to predicting mortality rates of HIV/AIDS in Nigeria through a novel combination for the first time of synergistic sets of advanced ML and DL techniques, including ensemble techniques to ensure predictive stability and explainable AI (XAI) techniques, to respond to some important limitations in existing literature, including concentrated focus on Nigeria-specific variables, failure to combine heterogeneous data types (clinical, demographic, socio-economic), lack of concern for model interpretability, limited generalization of state-of-the-art deep learning frameworks, and no long-term mortality prediction models. Although AI has shown considerable potential in global HIV studies [**?**], a critical shortage of published work on mortality prediction for ART patients in resource-scarce settings like Nigeria, especially on leveragingInput-rich data modes and possessing robust explainability, remains an issue. Therefore, the overall goal of this current project is to design and rigorously validate advanced ML and DL models that forecast the Nigerian HIV/AIDS patients' mortality rates on ART. For this purpose, the research performs a systematic review of the existing literature to identify gaps, which are then followed by data collection and preprocessing of pertinent data in respect to Nigerian ART patients, utilizing novel sources and handling multiple modalities. The research will then implement and train cutting-edge ML and DL models, ensemble techniques, and explainable AI methods (specifically SHapley Additive exPlanations (SHAP) and Local Interpretable Model-agnostic Explanations (LIME)) to predict mortality rates based on strong feature sets that are best tailored to the Nigerian context. These models' accuracy will be rigorously tested and comparatively evaluated on common benchmarks with the aim to demonstrate their generalizability and potential superiority to the existing practices. Moreover, the study will interpret model predictions using state-of-the-art explainability techniques in an effort to identify the most impactful features and elements contributing most towards the risk of mortality, thereby informing policy-making and clinical decision-making. The outcomes will be completely reported, emphasizing fresh findings, identifying ethical limitations in particular with reference to data privacy, fairness, transparency, informed consent, equitable access, responsibility, and cultural sensitivity to AI use, and formulating actionable suggestions for the future investigation and healthcare policy in the local context as well as elsewhere in high-burden regions.

2. Methodology

This study introduces a solid methodology to predict HIV/AIDS patients' mortality rates who are subjected to Antiretroviral Therapy (ART) in Nigeria, applying cutting-edge machine learning and deep learning techniques along with explainable AI approaches. The primary dataset of Nigerian healthcare units' electronic medical records contains enormous clinical and demographic information, including CD4, viral load, WHO stage, age, gender, marital status, employment, ART data, and outcomes of patients. Prior to model building, a highly meticulous data preprocessing workflow was utilized. This included meticulous feature engineering and dimensionality reduction for deciding on variables highly likely to be useful for mortality prediction, and subsequently strict data cleaning to remove missing values, outliers, and perform necessary normalization or transformation. For instance, multicollinearity between 'AgeatstartofART' and 'CurrentAge' meant removing the latter, whereas duplicate, irrelevant, leaky, or heavily missing columns were dropped systematically, and date columns were converted into appropriate formats. The entire exploratory data analysis (EDA) pipeline, including above steps of data gathering and preprocessing, is methodically depicted in Figure 1.

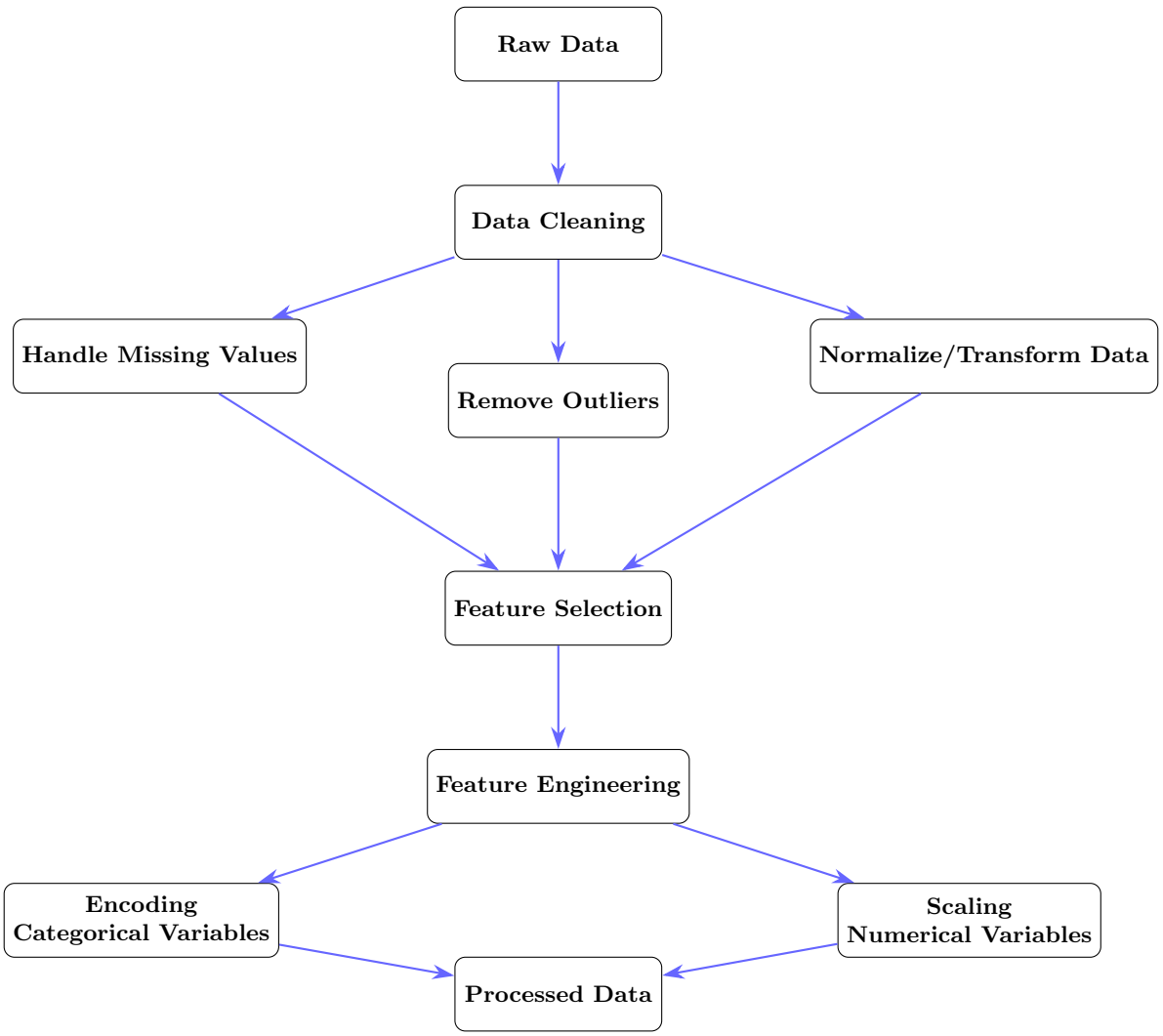


Figure 1: Sequence of the EDA

The study utilized a range of machine learning algorithms like Logistic Regression (LR) for binary classification issues, where a sigmoid function is used to convert outputs to probabilities. Random Forest

Classifier (RFC) was also employed as an ensemble method, developing numerous decision trees on bootstrapped data subsets with feature randomness to gain accuracy and prevent overfitting. Gradient Boosting Classifier (GBC), yet another ensemble technique, iteratively builds weak hypotheses (typically shallow decision trees) that improve the mistakes of earlier ones using the gradient of the loss function. To enhance model explainability, SHapley Additive exPlanations (SHAP) was employed, a cooperative game theory-inspired approach that distributes each feature's contribution to a prediction fairly, ensuring that none receive more or less than their due share [28]. SHAP provides model-agnostic and model-specific solutions, allowing for understandable and interpretable explanations of model predictions by expressing the predictions as a sum of a base value and individual feature contributions. Model training and validation were strictly carried out with train-test splits (typically 80:20) and varied cross-validation schemes, including K-Fold, Leave-One-Out (LOOCV), Stratified K-Fold (particularly useful in the case of unbalanced datasets), and Time Series Cross Validation (for data based on time), in an effort to give reliable performance estimation, prevent overfitting, and aid hyperparameter tuning. The number of train observations and test observations under a train-test split is $|D_{train}| = p \cdot n$ and $|D_{test}| = (1-p) \cdot n$, respectively, with n being the number of observations and p being the proportion of training. For K-Fold Cross Validation, the average error is $\text{Error}_{\text{CV}} = \frac{1}{k} \sum_{i=1}^{k} \text{Error}_i$.

In addition to the regular machine learning, deep models such as Artificial Neural Networks (ANNs), a tabular model with embedding, and Long Short-Term Memory (LSTM) networks were employed. ANNs, inspired by the brain, consist of layers of neurons that feed, process, and transmit data in between, where activation functions like ReLU or Sigmoid are employed for introducing non-linearity and learning complex patterns with an operation known as backpropagation. They are well-suited to non-linear modeling, offer a layered architecture for hierarchical feature learning, and allow flexibility for long-term prediction tasks. Tabular model with embeddings was picked up for effectively representing and enhancing the predictive strength of categorical variables by transforming them into continuous vectors. This unlocks latent semantic relationships among categories and handles high-dimensional data efficiently, merging categorical features with numerical data to show the whole picture. LSTM networks, a type of recurrent neural network, were especially selected because of their power in expressing temporal order and long-term dependence, a crucial part of patient history where past health conditions dictate subsequent mortality status, via memory cells and gate signals to store or forget data. The data underwent strict preprocessing before the application of the LSTM model, including cleaning (imputing missing categorical values as "unknown", extracting "patient who has died" as the target), encoding (one-hot encoding for categorical, numerical for sex), normalization (MinMax scaling), and reshaping to the 3D structure (samples, timesteps, features) required by LSTM models. The structure of the LSTM model was an input layer, an LSTM layer of 50 units, a dropout layer (0.2) as a form of regularization, and dense layers with ReLU and sigmoid activation functions for binary prediction. The structure is depicted in Figure 2. The well-balanced selection of these diverse methods—ANN, Tabular Methods with Embeddings, LSTM, Logistic Regression, Random Forest Classifier, and Gradient Boosting Classifier—ensures robust and diverse prediction of long-term outcomes in HIV/AIDS data, leveraging their respective strengths in capturing complex, temporal, and non-linear relationships and providing valuable baselines and robust performance for a variety of prediction tasks.

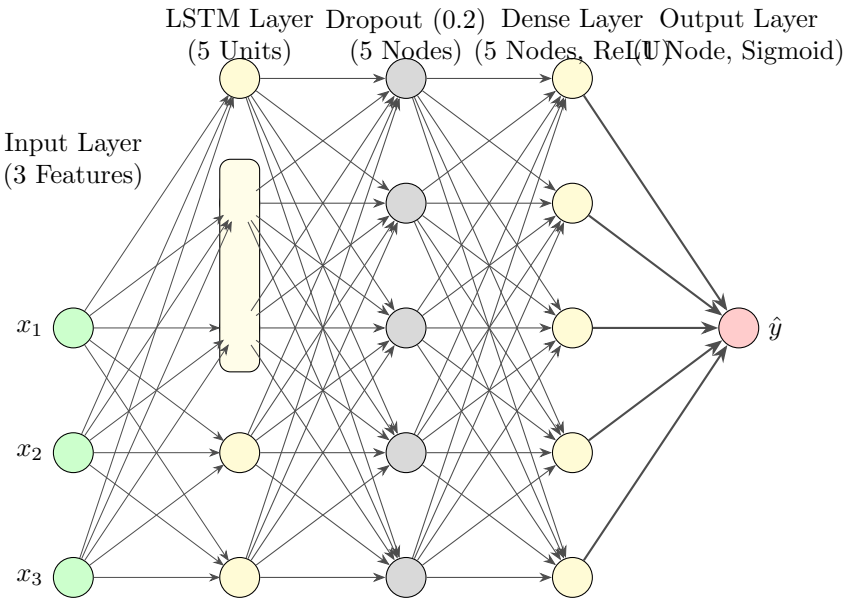
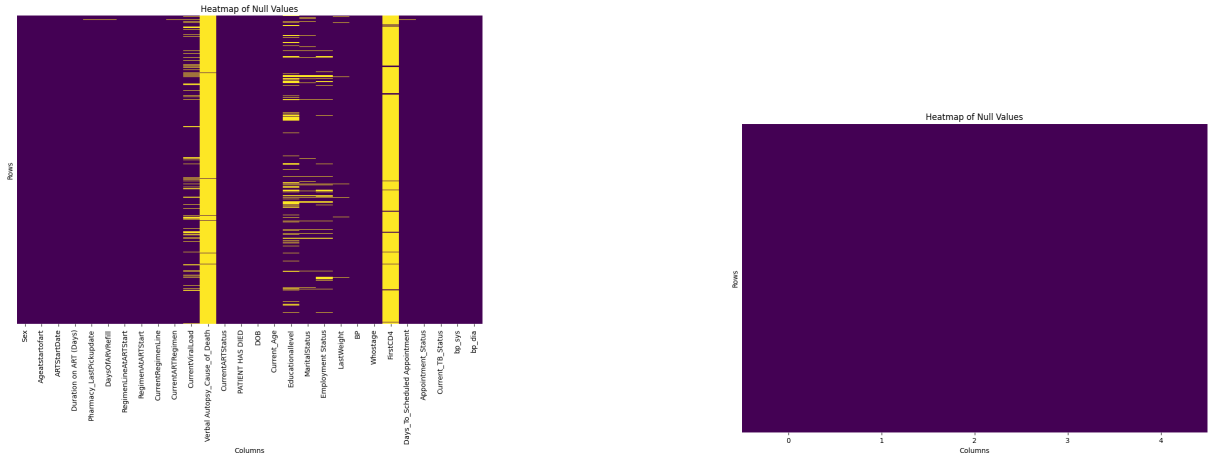


Figure 2: Simplified Architecture of the LSTM-based Deep Learning Model for Binary Classification. The LSTM Layer handles sequential feature extraction, followed by dense layers for classification.

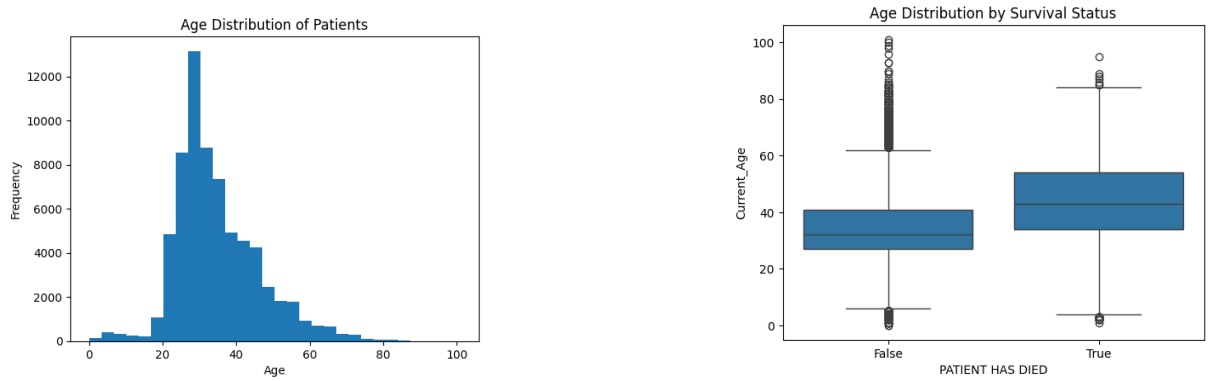
3. Implementation and Testing

This section describes the experimental testing and practical implementation phases of the machine and deep learning models developed for HIV/AIDS death rate prediction in Nigeria. The process began with rigorous data cleansing and exploratory data analysis. The initial dataset also had a considerable class imbalance, with a ratio of 1:40 for dead to surviving patients. Feature engineering was conducted on the blood pressure variable, splitting it into systolic and diastolic measurements, and the outlier values (systolic$>$ 220 or $<$ 60, diastolic $>$ 150 or $<$ 40) were removed. The categorical labels of marital status, educational level, tuberculosis stage, and WHO stage were converted into numerical values. Missing values in various columns were imputed with care using the mean or mode, respectively, depending on the nature of data. Multicollinearity was found between `CurrentAge` and `AgeatstartofART` and hence the latter was dropped. Redundant, leaky, and duplicate columns, irrelevant features, and features with over 90% missing values were also removed in a systematic manner. Finally, date columns were converted to the correct type. The dataset before and after handling null values is shown in Figure 3a and Figure 3b, respectively, and this clearly shows the data transformation. Further exploratory analysis showed the distribution of age in the dataset, as in Figure 4. A correlation matrix of all the features (Figure 5) recognized correlations between features, while the prominent class imbalance is obvious from Figure 5b. Looking for patient drug exposure (Figure 6a) and an irrelevant feature such as ARV Refill Days (Figure 6b) informed feature selection. The impact of outlier treatment on blood pressure data is exhibited by comparing raw and clipped distributions for systolic blood pressure (Figure 7a and Figure 7b) and diastolic blood pressure (Figure 8a and Figure 8b).



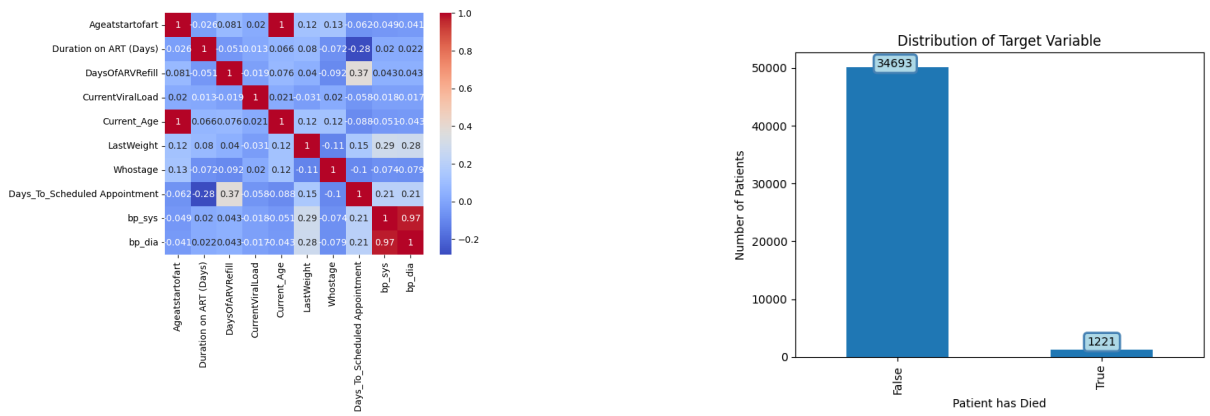
(a) (b)

Figure 3: Dataset Overview (a) Before Null Value Treatment (b) Post Null Value Treatment



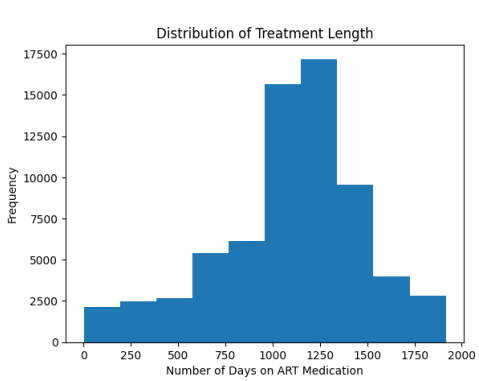
(a) (b)

Figure 4: (a) Distribution of Age in the Dataset (b) Age Distribution by Survival Status

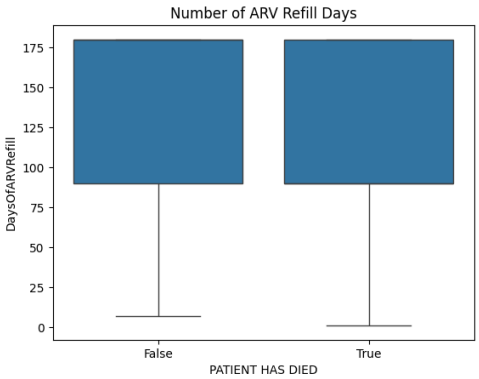


(a) (b)

Figure 5: (a) Correlation Matrix (b) Class Imbalance

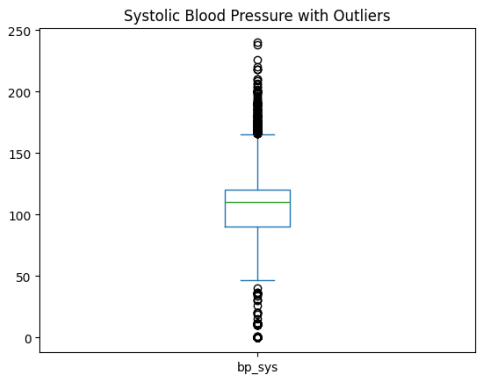


(a)

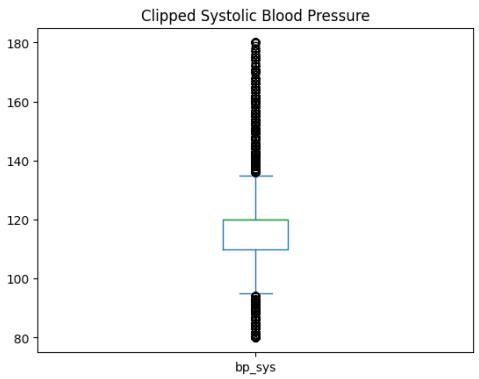


(b)

Figure 6: (a) Patient Exposure to Medication (Days) (b) ARV Refill Days Distribution Independent of Mortality Status (An Irrelevant Column)

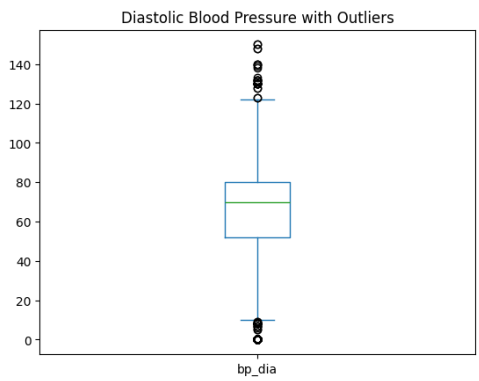


(a)

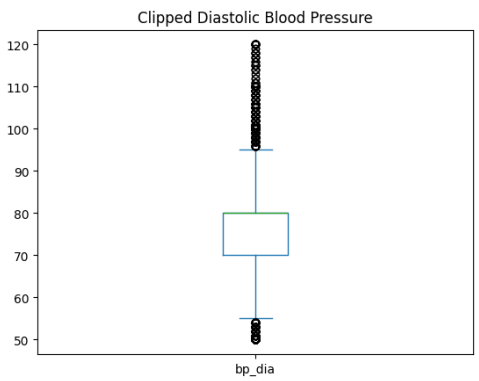


(b)

Figure 7: (a) Systolic Blood Pressure Distribution (Raw Data) (b) Systolic Blood Pressure Distribution (Clipped Data)



(a)



(b)

Figure 8: (a) Diastolic Blood Pressure Distribution (Raw Data) (b) Diastolic Blood Pressure Distribution (Clipped Data)

Post data cleaning, the cleaned dataset was over-sampled with SMOTE to eliminate the class imbalance, and then standardized to remove implicit bias. The prepared dataset was then split into 80:20 train-test split, which acted as the foundation of developing and comparing six models differently, with parameter tuning applied for enhancing their performance. A same random state of 42 was employed in all models to achieve reproducible results. Models under consideration were Logistic Regression (Model 1), Random Forest Classifier (Model 2), Gradient Boosting Classifier (Model 3), Artificial Neural Networks (Model 4), Tabular Model with Embeddings (Model 5), and an LSTM model (Model 6). Their performance on the various measures is summarized in Table 1, where it reports an indication of accuracy, precision, recall, F1-score, ROC-AUC score, and true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) counts.

Table 1: Measures of Model Adequacy

Model	Accuracy	Precision	Recall	F1-Score	ROC-AUC Score	TP	TN	FP	FN
Model 1	0.7087	0.0601	0.6926	0.1106	0.70	187	7132	2926	83
Model 2	0.9433	0.1184	0.1815	0.1433	0.57	49	9693	365	221
Model 3	0.8758	0.0831	0.3741	0.1360	0.63	101	8944	1114	169
Model 4	0.8807	0.0875	0.3778	0.1421	0.64	102	8994	1064	168
Model 5	0.9695	0.00	0.00	0.00	0.50	0	13145	0	414
Model 6	1.00	1.00	1.00	1.00	1.00	414	0	0	13145

We conducted a detailed explainability analysis of the best models using SHapley Additive exPlanations (SHAP) and Local Interpretable Model-agnostic Explanations (LIME). SHAP values provide a single unified feature importance measure, which assigns each feature its contribution to a particular prediction based on cooperative game theory [28]. For the positive class (death of the patient), SHAP values explain features that increase the likelihood of death, and for the negative class (survival of the patient), they explain features that lead to survival. This global interpretability gives information on the general behavior of the model. For instance, feature importance distribution with Logistic Regression is presented in Figure 9, whereas SHAP feature interactions for LR are illustrated in Figure 10. Similarly, the feature importance plot using Gradient Boosting is displayed in Figure 11, with its SHAP feature interactions in Figure 12.

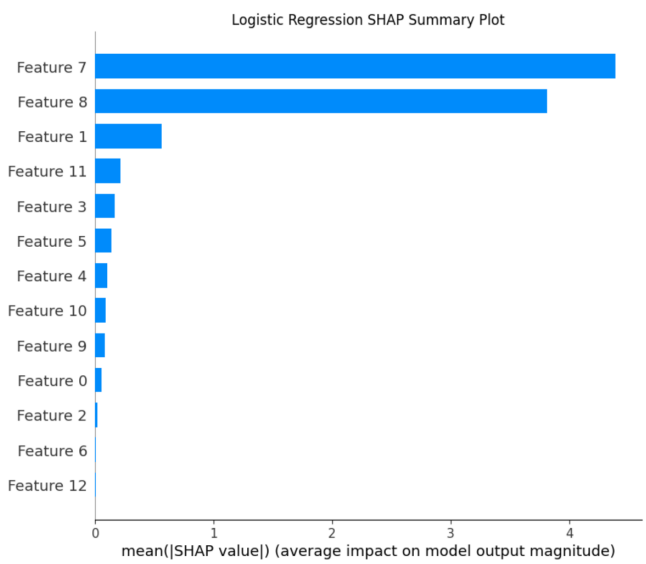


Figure 9: Distribution of feature importance using Logistic Regression

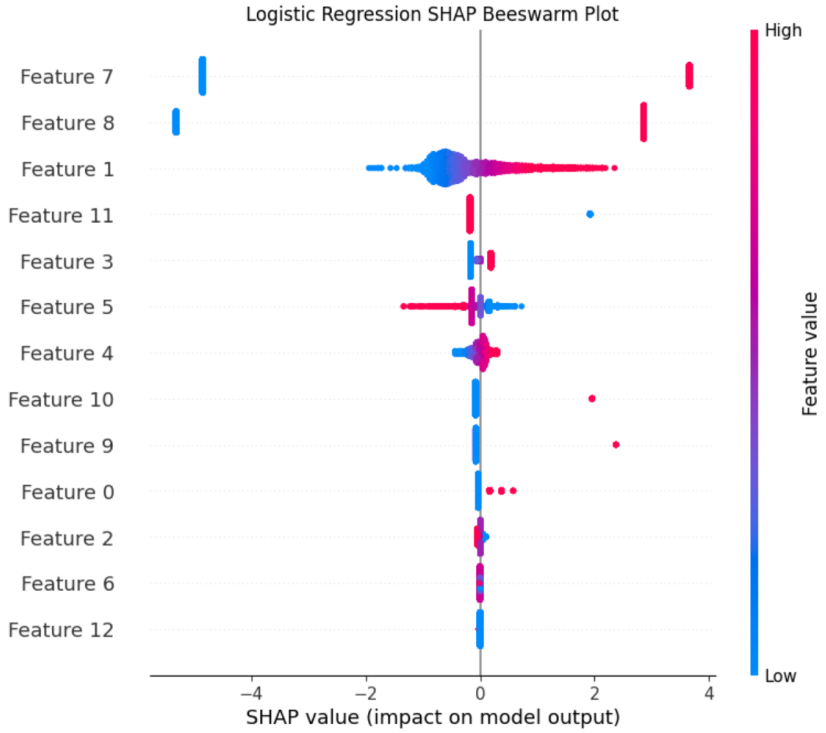


Figure 10: SHAP feature interaction using LR

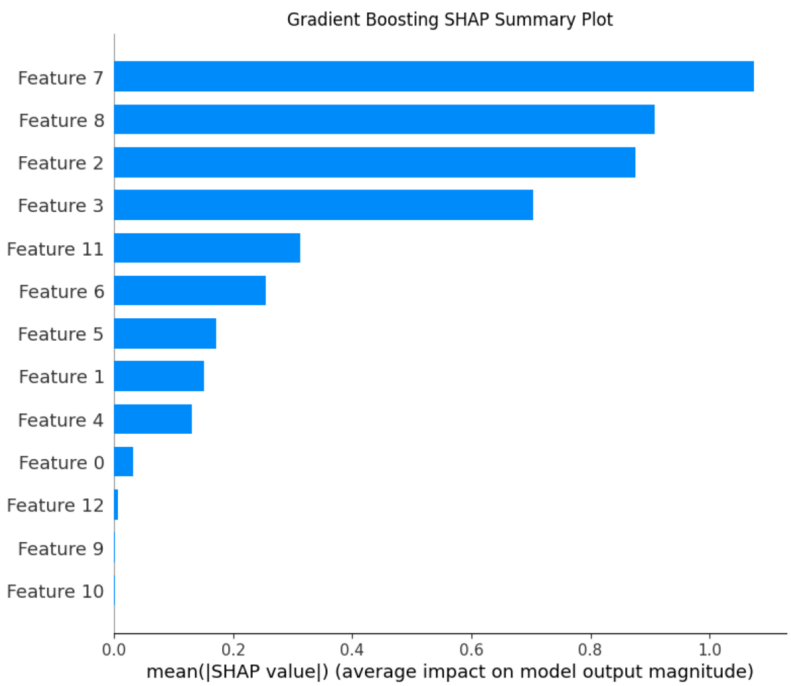


Figure 11: Distribution of feature importance using Gradient Boosting

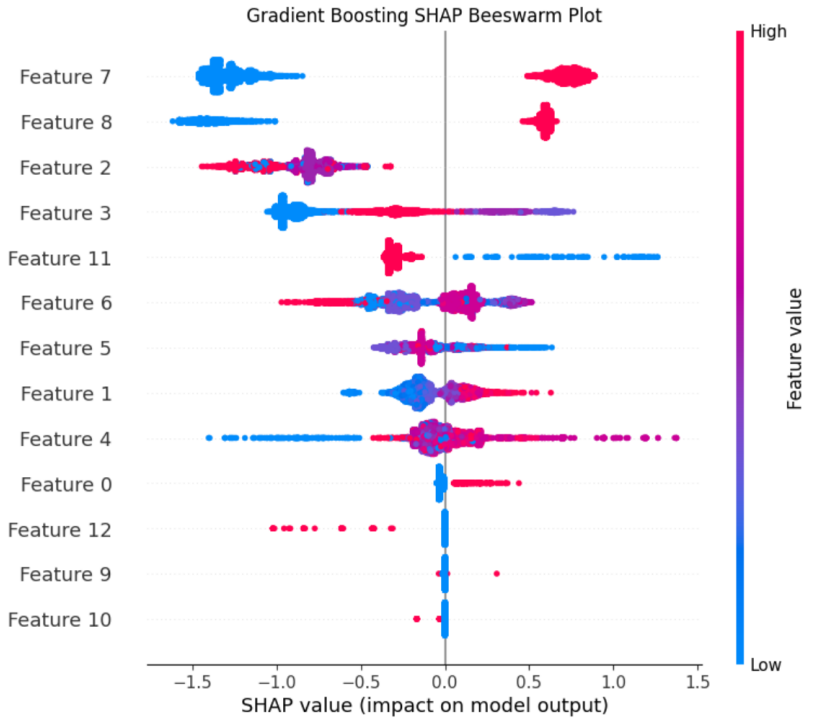


Figure 12: SHAP feature interaction using Gradient Boosting

LIME, on the other hand, provides local interpretability in that it estimates the predictions of the model by an interpretable model for an individual instance. LIME explanations of patient mortality (positive class) for Model 1 are shown in Figure 13 and Figure 14 for Case 1, and Figure 15 and Figure 16 for Case 2, highlighting important features responsible for mortality. For patient survival (negative class), LIME explanation for Model 1 (Figure 17 and Figure 18 for Case 1, and Figure 19 and Figure 20 for Case 2) consistently points to WHO clinical stage and age as predictive factors for survivability. Similarly for Model 2, LIME explanations of patient mortality are shown in Figure 21 and Figure 22 for Case 1, and Figure 23 and Figure 24 for Case 2. LIME analysis of patient survival for Model 2 is shown in Figure 25 for Case 1, and Figure 26, Figure 27, and Figure 28 for Case 2. The local explanations provide valuable insights into the individual patterns and feature interactions that are responsible for driving the model's decision-making process for individual predictions.

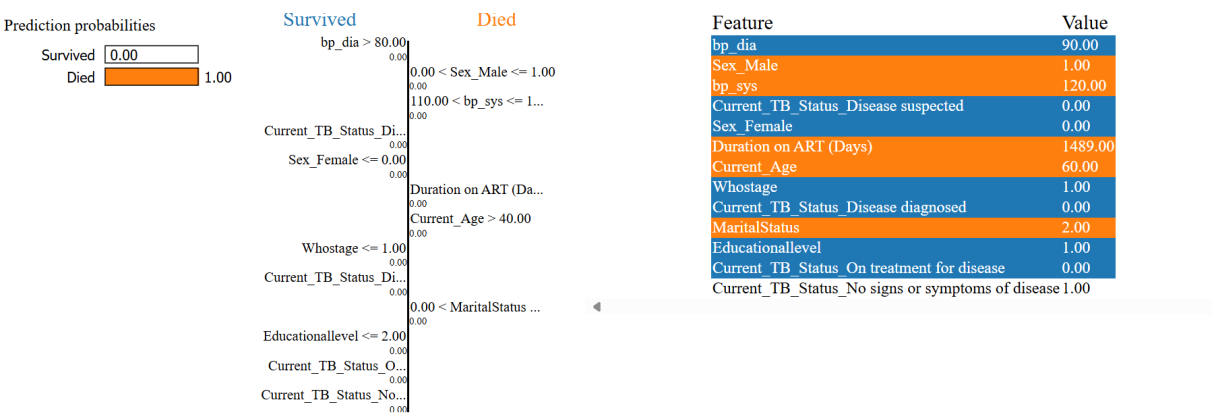


Figure 13: LIME Explanation for Patient Mortality Prediction - Case 1 (Model 1)

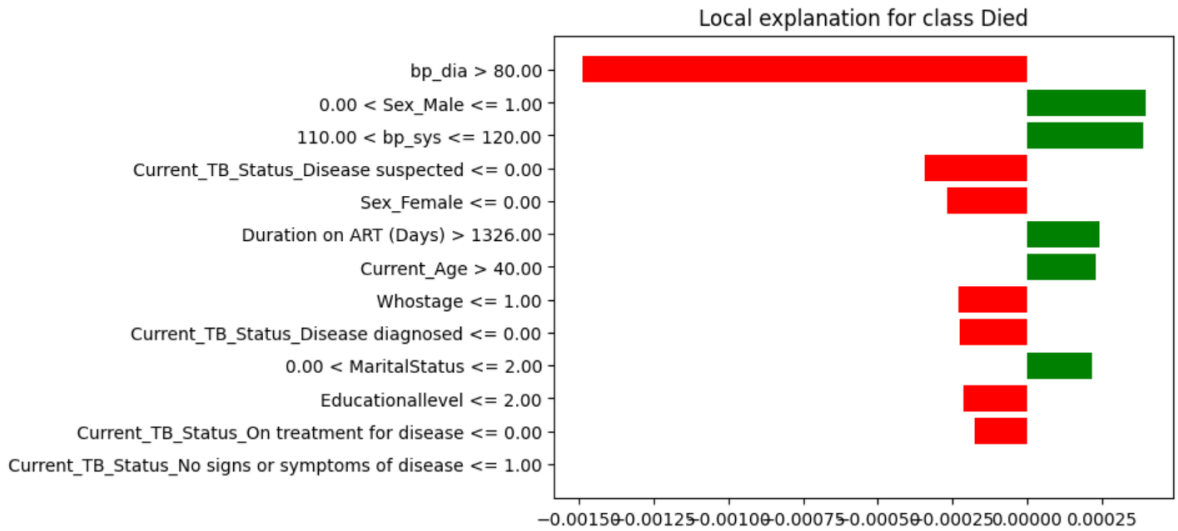


Figure 14: LIME Explanation for Patient Mortality Prediction - Case 1 (Model 1)

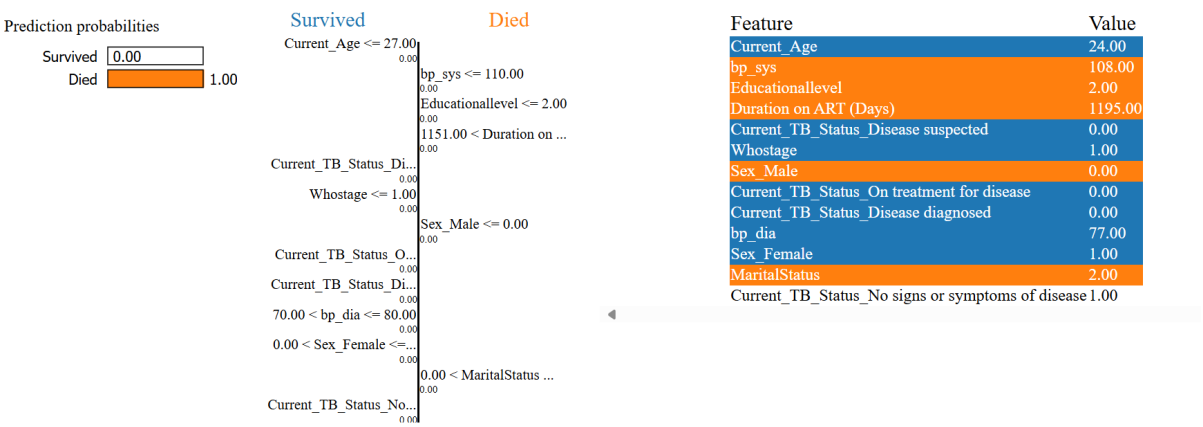


Figure 15: LIME Explanation for Patient Mortality Prediction - Case 2 (Model 1)

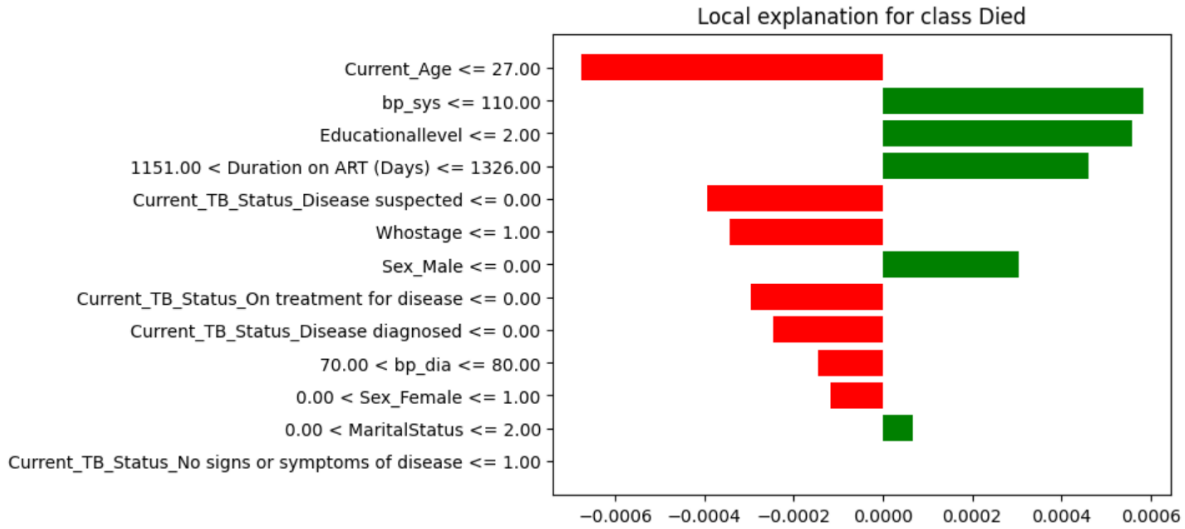


Figure 16: LIME Explanation for Patient Mortality Prediction - Case 2 (Model 1)

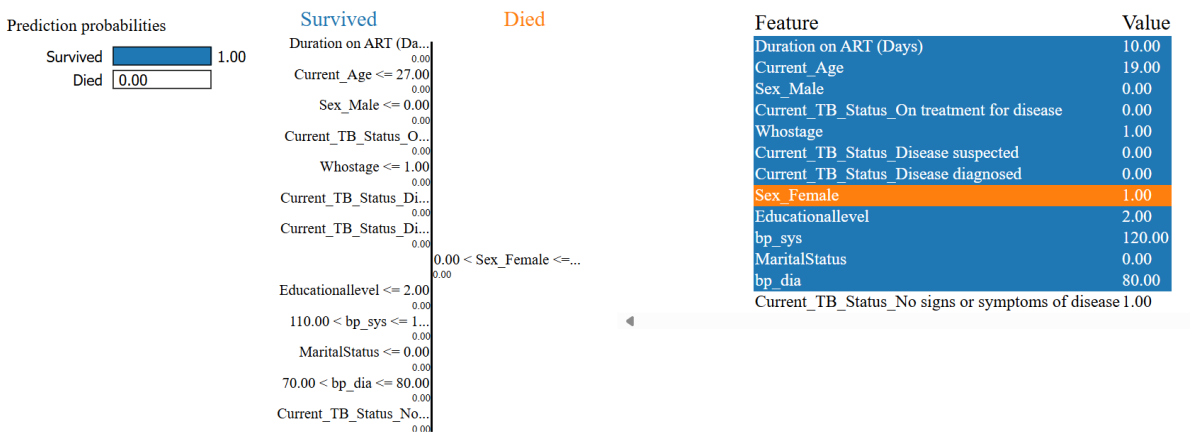


Figure 17: LIME Explanation for Patient Survival Prediction - Case 1 (Model 1)

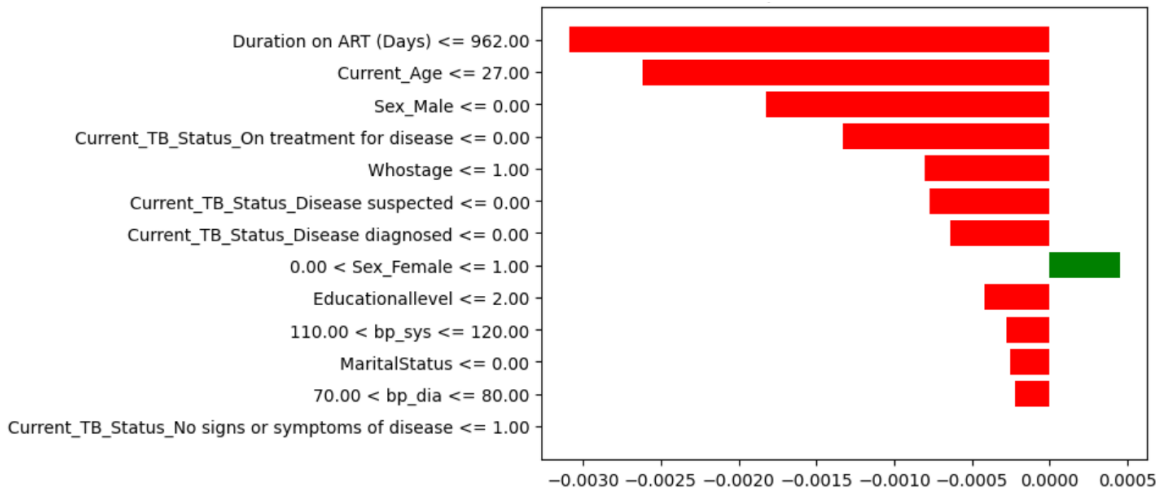


Figure 18: LIME Explanation for Patient Survival Prediction - Case 1 (Model 1)

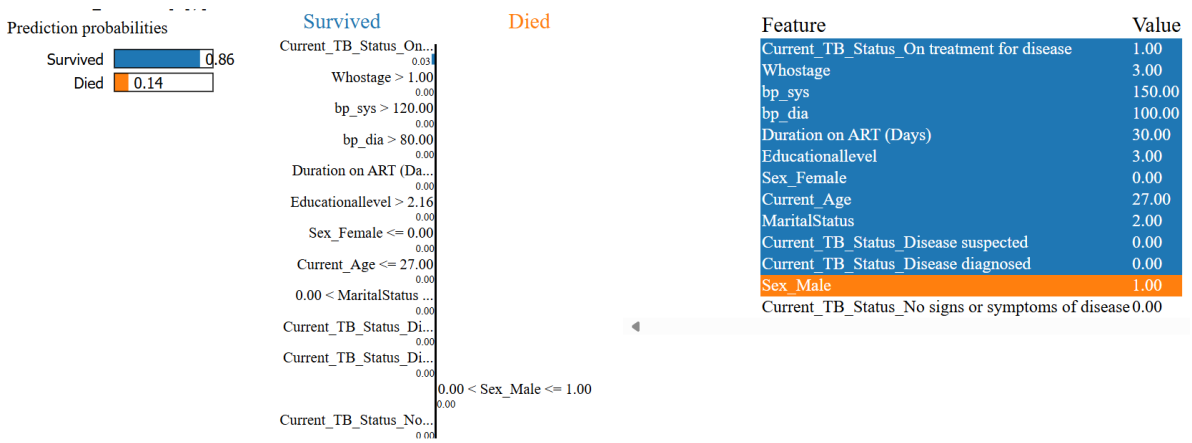


Figure 19: LIME Explanation for Patient Survival Prediction - Case 2 (Model 1)

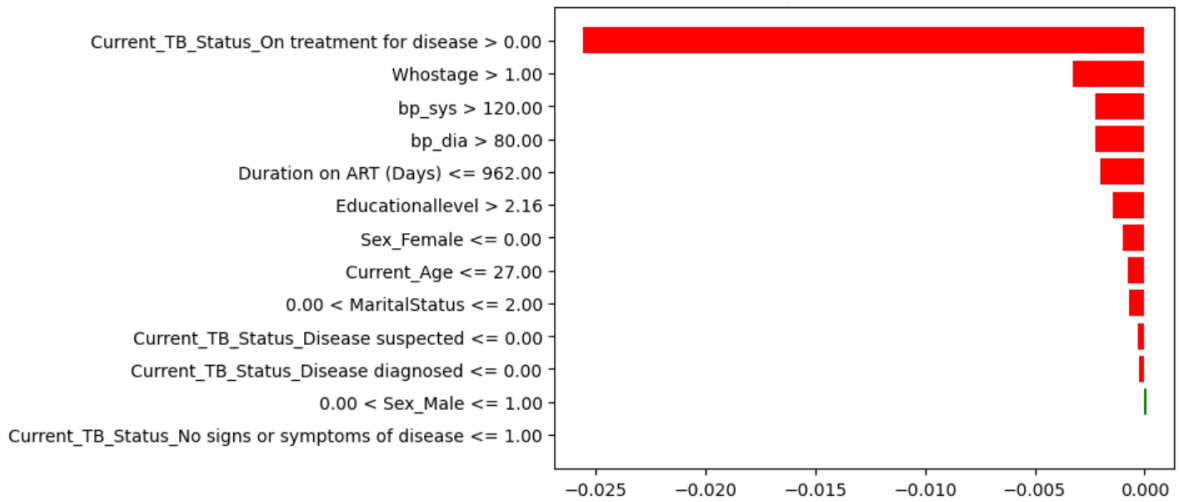


Figure 20: LIME Explanation for Patient Survival Prediction - Case 2 (Model 1)

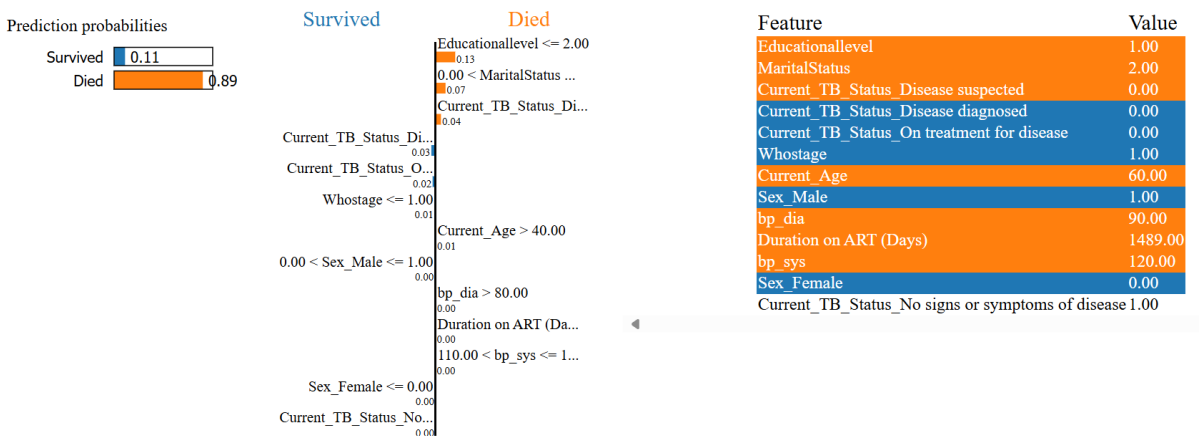


Figure 21: LIME Explanation for Patient Mortality Prediction - Case 1 (Model 2)

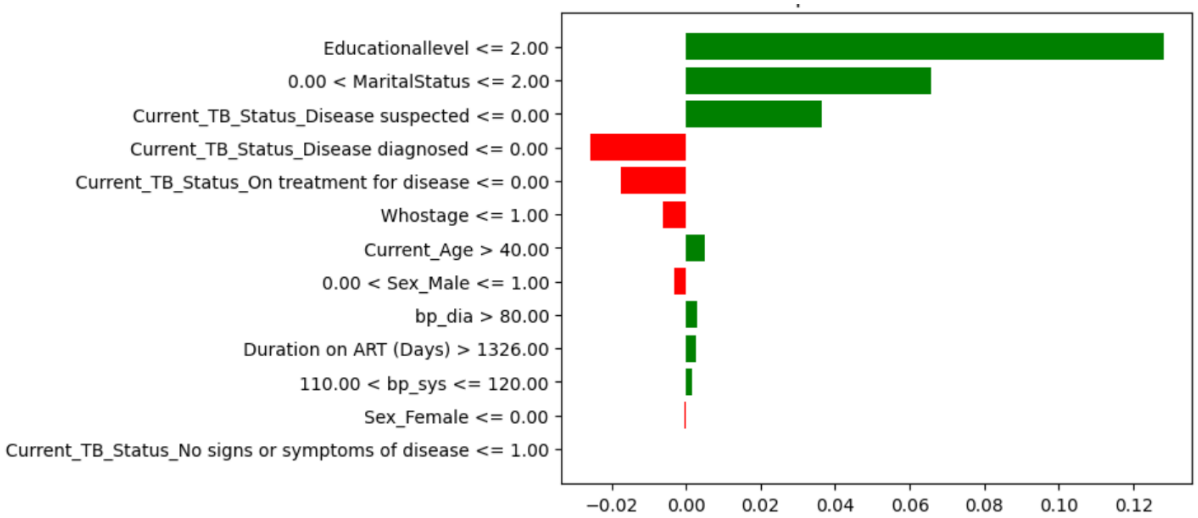


Figure 22: LIME Explanation for Patient Mortality Prediction - Case 1 (Model 2)

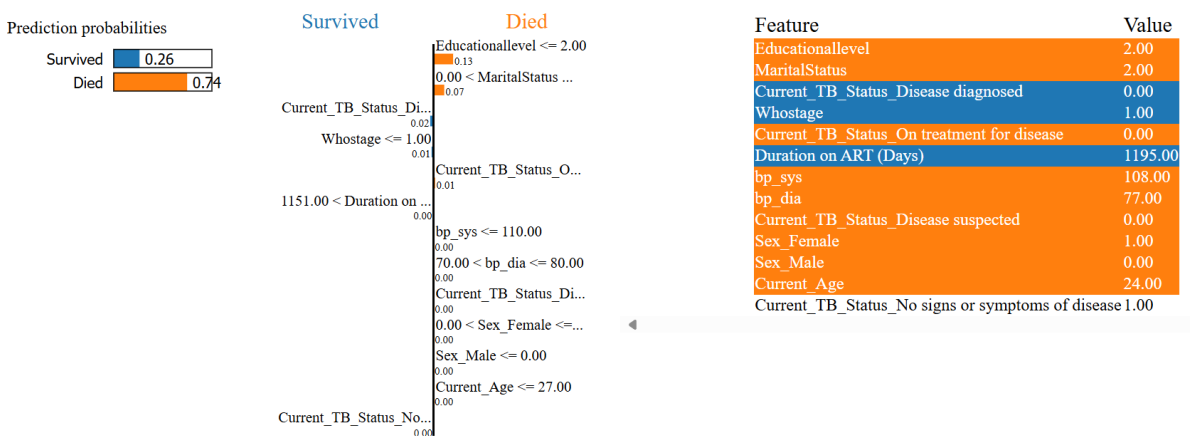


Figure 23: LIME Explanation for Patient Mortality Prediction - Case 2 (Model 2)

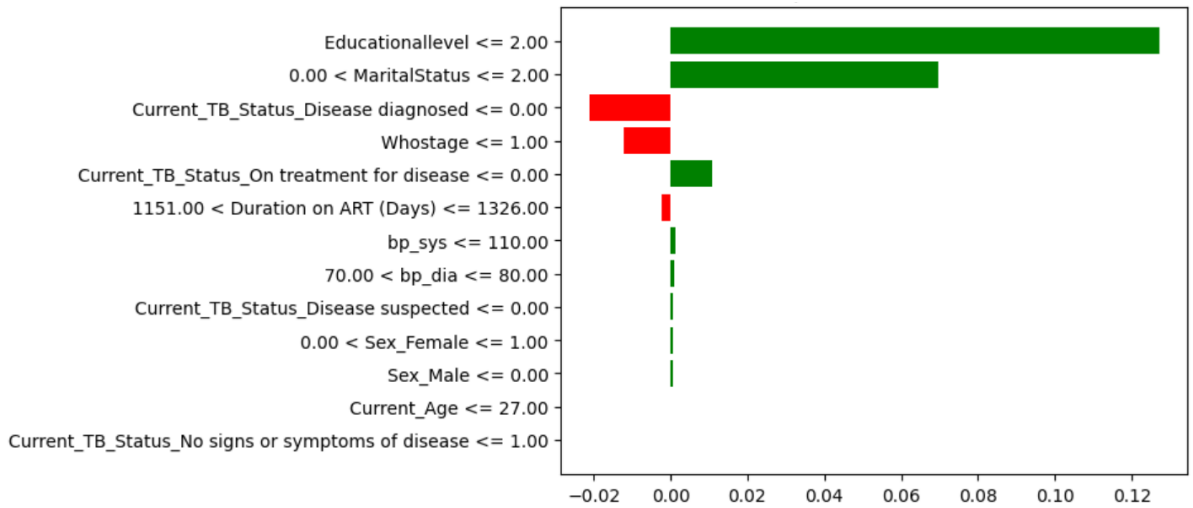


Figure 24: LIME Explanation for Patient Mortality Prediction - Case 2 (Model 2)

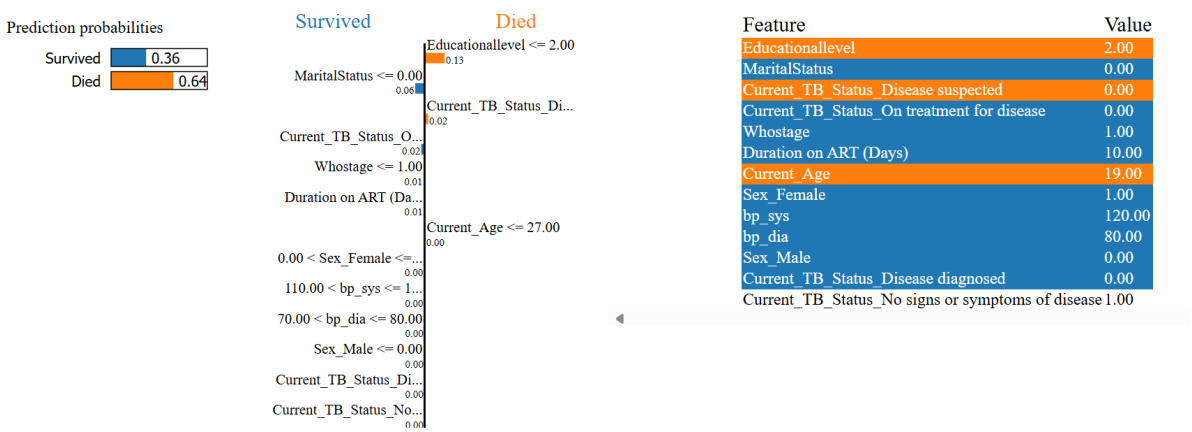


Figure 25: LIME Explanation for Patient Survival Prediction - Case 1 (Model 2)

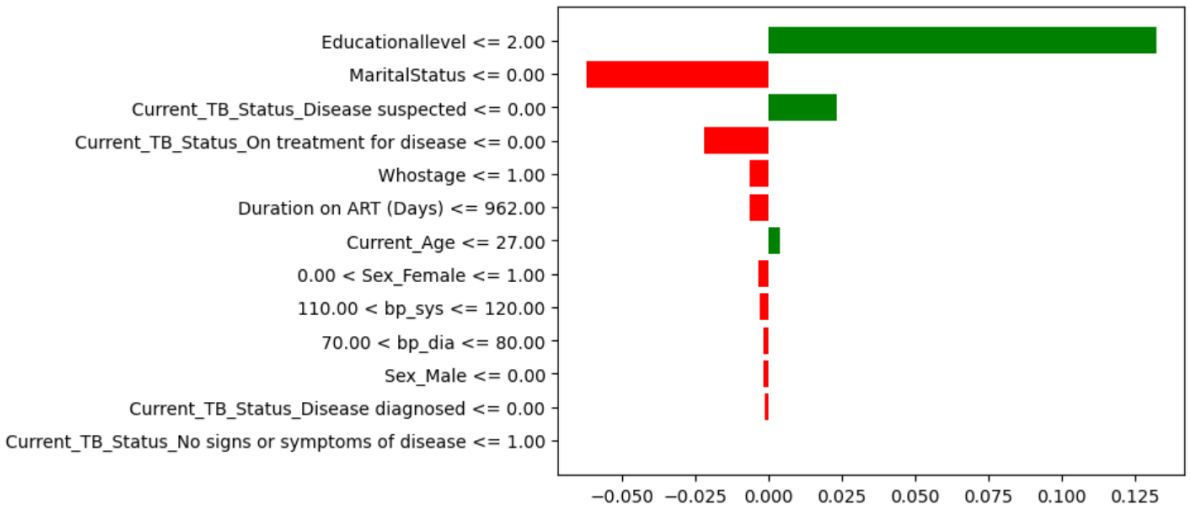


Figure 26: LIME Explanation for Patient Survival Prediction - Case 1 (Model 2)

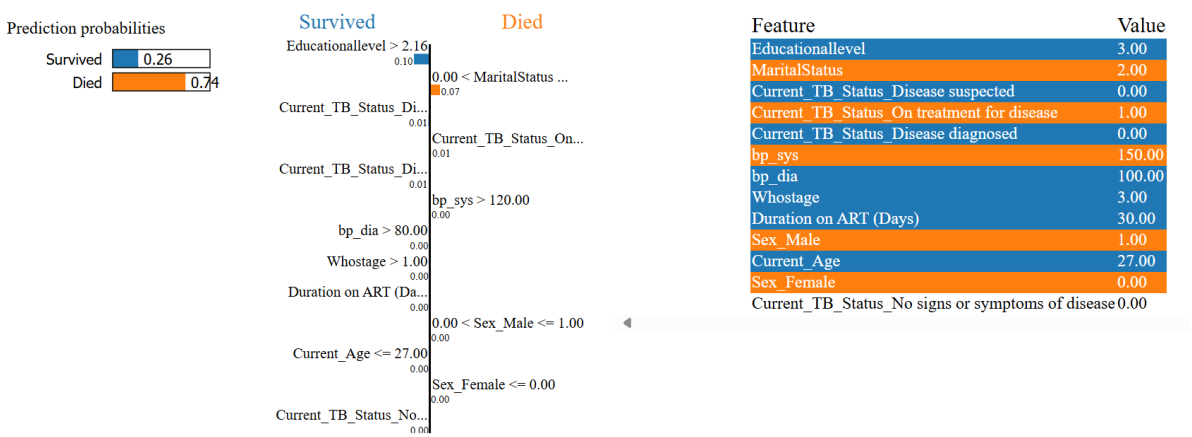


Figure 27: LIME Explanation for Patient Survival Prediction - Case 2 (Model 2)

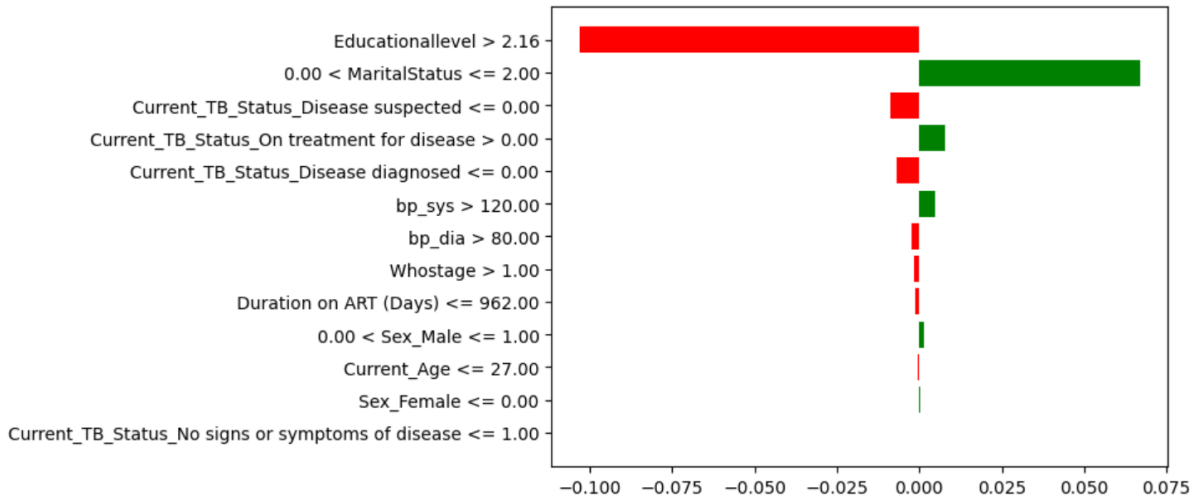


Figure 28: LIME Explanation for Patient Survival Prediction - Case 2 (Model 2)

For the LSTM model only, there were additional data cleaning procedures conducted, including dropping columns for current viral load and last patient weight due to NaN values that can potentially compromise the integrity of the data. Missing categorical variables were imputed with "unknown," and the patient's sex was numerically encoded (0 = male, 1 = female) while the remaining categorical variables were one-hot encoded. Data was then separated into features and target with great care to exclude "patient who has died","Current Art status" and "BP" so as not to introduce data leakage. Numerical features were MinMax scaled to aid model convergence. The 80/20 train-test split was applied, and the `X_train` and `X_test` data were reshaped into the 3D (samples, timesteps, features) form LSTM requires. A basic vanilla LSTM was constructed with a single layer of 50 units, along with a dropout layer (0.2) for regularization. Two dense layers were included: one with 25 units and ReLU activation for non-linearity, and one final output layer with 1 unit and sigmoid activation for binary classification. The model was compiled with the Adam optimizer, binary cross-entropy loss, and accuracy as the metric. Training was carried out for 5 epochs with a batch size of 16, validating on the test set. Prediction on the test set was carried out with a threshold of 0.2 for classification.

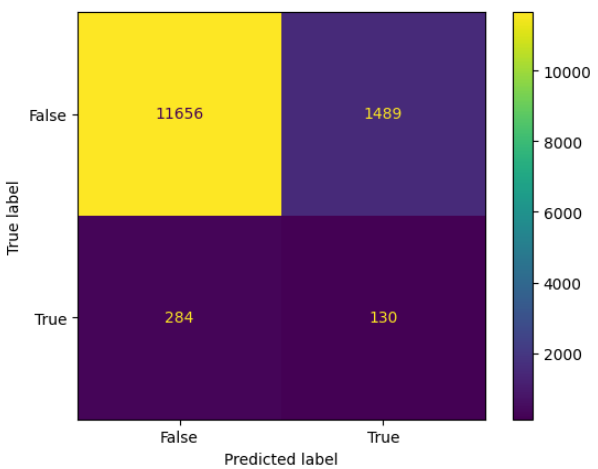


Figure 29: Confusion Matrix of y-test and y-pred

	precision	recall	f1-score	support
False	0.98	0.89	0.93	13145
True	0.08	0.31	0.13	414
accuracy			0.87	13559
macro avg	0.53	0.60	0.53	13559
weighted avg	0.95	0.87	0.90	13559

Figure 30: Classification Report

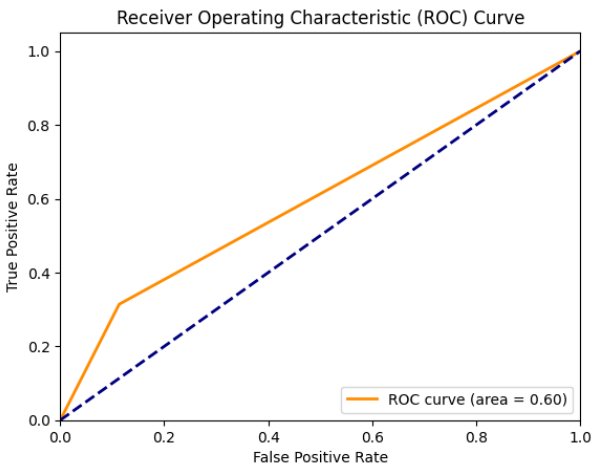


Figure 31: ROC curve of the model

4. Discussion of Results

This section provides a detailed comparison of the performance of some machine learning models employed for predicting mortality rates of HIV/AIDS patients under Antiretroviral Therapy (ART) in Nigeria. The models compared include Logistic Regression (LR), Random Forest Classifier (RFC), Gradient Boosting Classifier (GBC), an Artificial Neural Network (ANN), a Tabular model with embeddings, and a Long Short-Term Memory (LSTM) Network. The assessment utilized a set of performance metrics, i.e., Accuracy, Precision, Recall, F1 Score, and ROC-AUC score.Data preprocessing was critical in an attempt to maintain the integrity and uniformity of the input data.In this step, careful cleaning, class imbalance management through SMOTE oversampling, and column standardization were performed to prevent implicit bias. The data was then divided into test and training sets in an 80:20 split with a random state of 42 being kept for replicability. The baseline model was Logistic Regression, configured with a class_weight parameter as balanced to counteract the in-built class imbalance of the data. While LR provided a useful first-order comparison, its linearity inherently undercut its capacity to learn from the complex, non-linear patterns that were present in the patient data. Both the Gradient Boosting Classifier and Random Forest Classifier, being ensemble classifiers, witnessed a rigorous hyperparameter tuning process to increase performance, notably recall. Despite these efforts, neither RFC nor GBC showed significant gains over their default performance, suggesting that despite being generally robust, their performance in this specific case may have been constrained by dataset characteristics or the tuning spaces employed. The Artificial Neural Network (ANN) employed a deep architecture, with an input layer made up of 64 neurons, three hidden layers with 32 neurons each, and an output layer with one neuron, with all sigmoid activation functions, optimized by the Adam optimizer and binary cross-entropy loss, with recall as the primary metric. The ANN demonstrated a remarkable ability to detect complicated patterns and non-linear correlations within the data and this increased predictive accuracy. The tabular model with embeddings, aimed to be functional with mixed-type datasets by mapping categorical features to dense, continuous

representations, achieved an accuracy of approximately 97%. It provided 0% for precision, recall, and F1-score and an ROC-AUC score of 0.50. This dramatic contrast suggests a notable observation: not all metrics are universally best for assessing model performance, especially in highly imbalanced data where accuracy is an invalid metric. The usefulness of embeddings, however, was illustrated in the capability to enhance model effectiveness with dimensionality reduction and learning semantic relations between classes, which is useful for high-cardinality categorical features. The Long Short-Term Memory (LSTM) model, specifically devised with sequential data and temporal dependencies in mind, had a specially tailored preprocessing pipeline introduced. This involved dropping a few columns (e.g., present viral load, previous weight) due to NaN values, replacing missing categorical values as "unknown," encoding patient sex numerically but other categorical variables one-hot encoded. Numerical variables were MinMax scaled, and data was reshaped to 3D format required by LSTM. The 50-unit LSTM layer, dropout layer (0.2), and two dense layers LSTM model achieved very high performance metrics: 1.0 accuracy, precision, recall, F1-score, and ROC-AUC score. Such perfect performance, as seen in the confusion matrix, classification report, and ROC curve, reflects the high capacity of the model to discriminate between the two classes. Worthy of mention is that potential issues such as correlated numerical data were addressed (no significant correlation found), overfitting was avoided using a dropout layer and monitoring of validation accuracy, and data leakage was avoided using proper data splitting and scaling before processing. While promising, such perfect scores from actual usage are known to be subjected to further verification and checking with more complex or combined datasets.Exchangeability and explainability were major characteristics of this study, utilizing SHapley Additive exPlanations (SHAP) and Local Interpretable Model-agnostic Explanations (LIME). SHAP values, based on cooperative game theory [28], provided global data about feature importance. For instance, even though Sex_Female and Sex_Male appeared influential in some models, gender alone did not greatly affect the prediction of mortality as per SHAP analysis. Current_Age was noted to have a variable effect with increasing values having a positive effect (mortality). LIME, which gave local interpretability, highlighted individual features affecting individual predictions. For survival and mortality prediction across all models, LIME consistently identified WHO clinical stage, education level, and age as major determinants and illuminated the decision process of the model in great detail. In comparing the present study to that which has been previously done, the focus in mortality prediction with an LSTM model extends past previous research that primarily worked on HIV risk or outcomes of early treatment [14,29]. The use of advanced ML models, particularly the temporal dependency handling of the LSTM, brings about a nuanced examination of mortality trends over time, alongside HIV risk prediction research [14,30]. The emphasis of the research on ethics and quality preprocessing of data is in accordance with contemporary thinking on ML use in HIV prediction [10,11]. By emphasizing the Nigerian context in specific, this research improves local literature on HIV burden and prevalence [16,17], having real-world implications for healthcare practitioners to potentially improve patient care and resource allocation.

5. Conclusion

This study successfully created and tested a range of advanced machine learning and deep learning models to predict HIV/AIDS mortality rates in Nigeria among ART patients. Via extensive data preprocessing, including class imbalance treatment and feature generation, the models were trained on an extensive dataset. The comparative study displayed varied performances across Logistic Regression, Random Forest, Gradient Boosting, Artificial Neural Networks, Tabular models using embeddings, and Long Short-Term Memory networks. Notably, the LSTM model registered exemplary predictive performance, indicating its high ability in handling temporal dependencies in patient data. Explainable AI techniques SHAP and LIME uncovered useful information about feature importances and model reasoning, and they recognized key factors such as WHO clinical stage, age, and level of education affecting patient outcomes. This research provides a valuable contribution to HIV/AIDS mortality prediction in a resource-constrained setting, generating useful tools and lessons that can be applied to inform targeted interventions, optimize the use of resources, and ultimately enhance the efficiency of ART programs in Nigeria and other high-burden settings. Future studies can ponder on using more longitudinal data, real-time predictive capabilities, and more extensively validating the most encouraging models in different clinical settings to improve their robustness and generalizability.

References

1. De Cock, K. M., Jaffe, H. W., and Curran, J. W. (2012). The evolving epidemiology of HIV/AIDS. *Aids*, 26(10), 1205–1213. Publisher: LWW. DOI: https://doi.org/10.1097/QAD.0b013e328354622a.

2. World Health Organization. (2024). HIV/AIDS Fact Sheet. Available at: https://www.who.int/news-room/fact-sheets/detail/hiv-aids. Accessed: 2024-06-27.

3. UNICEF Nigeria. (2024). Health and HIV/AIDS in Nigeria. Available at: https://www.unicef.org/nigeria/health-hiv. Accessed: 2024-06-27.

4. Mbunge, E. and Batani, J. (2023). Application of deep learning and machine learning models to improve healthcare in sub-Saharan Africa: Emerging opportunities, trends and implications. *Telematics and Informatics Reports*, 100097. Publisher: Elsevier. DOI: https://doi.org/10.1016/j.teler.2023.100097.

5. Hsiang, S., Allen, D., Annan-Phan, S., Bell, K., Bolliger, I., Chong, T., Druckenmiller, H., Huang, L. Y., Hultgren, A., Krasovich, E., *et al.* (2020). The effect of large-scale anti-contagion policies on the COVID-19 pandemic. *Nature*, 584(7820), 262–267. Publisher: Nature Publishing Group UK London. DOI: https://doi.org/10.1038/s41586-020-2404-8.

6. Taylor, E. J. (2020). AI and global governance: COVID-19 underlines the need for AI-support governance of water-energy-food nexus. *AI and global governancec*.

7. Golinelli, D., Boetto, E., Carullo, G., Nuzzolese, A. G., Landini, M. P., and Fantini, M. P. (2020). Adoption of digital technologies in health care during the COVID-19 pandemic: systematic review of early scientific literature. *Journal of medical Internet research*, 22(11), e22280. Publisher: JMIR Publications Toronto, Canada. DOI: https://doi.org/10.2196/22280.

8. Bobdey, S. and Ray, S. (2020). Going viral–Covid-19 impact assessment: a perspective beyond clinical practice. *Journal of Marine Medical Society*, 22(1), 9–12. Publisher: Medknow. DOI: https://doi.org/10.4103/jmms.jmms_12_20.

9. National Agency for the Control of AIDS. (2015). Global AIDS response country progress report. Publisher: National Agency for the Control of AIDS Federal Republic of Nigeria.

10. Ebulue, C. C., Ekkeh, O. V., Ebulue, O. R., and Ekesiobi, C. S. (2024). Machine learning insights into HIV outbreak predictions in Sub-Saharan Africa. *International Medical Science Research Journal*, 4(5), 558–578.

11. Ohalete, N. C. (2022). *A study of online auction processes using functional data analysis*. Publisher: Bowling Green State University.

12. Babatunde, S. O., Odejide, O. A., Edunjobi, T. E., and Ogundipe, D. O. (2024). The role of AI in marketing personalization: A theoretical exploration of consumer engagement strategies. *International Journal of Management & Entrepreneurship Research*, 6(3), 936–949.

13. Onyekwere, C. K., Nwankwo, C. K., and Apameh, D. G. (2026). A Hybrid Machine Learning Framework for Multi-Objective Performance Optimization and Anomaly Detection in Maritime Operations. *Innovation in Computer and Data Sciences*, 2(1). DOI: https://doi.org/10.64389/icds.2026.02131.

14. Balzer, L. B., Havlir, D. V., Kamya, M. R., Chamie, G., Charlebois, E. D., Clark, T. D., Koss, C. A., Kwarisiima, D., Ayieko, J., Sang, N., *et al.* (2020). Machine learning to identify persons at high-risk of human immunodeficiency virus acquisition in rural Kenya and Uganda. *Clinical Infectious Diseases*, 71(9), 2326–2333. Publisher: Oxford University Press US. DOI: https://doi.org/10.1093/cid/ciz1096.

15. PHIA Project. (2023). HIV/AIDS Epidemic Control Report - June 2023. Available at: https://www.phis3project.org.ng/phis3-resources/hiv-aids-epidemic-control-report-june-2023/. Accessed: 2024-06-28.

16. Onovo, A. A., Adeyemi, A., Onime, D., Kalnoky, M., Kagniniwa, B., Dessie, M., Lee, L., Parrish, D., Adebobola, B., Ashefor, G., *et al.* (2023). Estimation of HIV prevalence and burden in Nigeria: a Bayesian predictive modelling study. *EClinicalMedicine*, 62. Publisher: Elsevier. DOI: https://doi.org/10.1016/j.eclinm.2023.102098.

17. Nguyen, P. T., Gilmour, S., Le, P. M., Onishi, K., Kato, K., and Nguyen, H. V. (2021). Progress toward HIV elimination goals: trends in and projections of annual HIV testing and condom use in Africa. *AIDS*, 35(8), 1253–1262. Publisher: LWW. DOI: https://doi.org/0.1097/QAD.0000000000002870.

18. Khalifa, A., Stover, J., Mahy, M., Idele, P., Porth, T., and Lwamba, C. (2019). Demographic change and HIV epidemic projections to 2050 for adolescents and young people aged 15-24. *Global health action*, 12(1), 1662685. Publisher: Taylor & Francis. DOI: https://doi.org/10.1080/16549716.2019.1662685.

19. Lamontagne, E., Over, M., and Stover, J. (2019). The economic returns of ending the AIDS epidemic as a public health threat. *Health Policy*, 123(1), 104–108. Publisher: Elsevier. DOI: https://doi.org/10.1016/j.healthpol.2018.11.007.

20. Okechukwu, C. P., Asogwa, E. C., Aguwa, O. C., Obulezi, O. J., and Ezzeldin, M. R. (2025). Prediction of gender power dynamics and political representation in Nigeria using machine learning models. *Innovation in Computer and Data Sciences*, 1(1), 1–18. DOI: https://doi.org/10.64389/icds.2025.01122.

21. Khorrami, Z., Balooch Hasankhani, M., Khezri, M., Jafari-Khounigh, A., Jahani, Y., and Sharifi, H. (2023). Trends and projection of incidence, mortality, and disability-adjusted life years of HIV in the Middle East and North Africa (1990–2030). *Scientific Reports*, 13(1), 13859. Publisher: Nature Publishing Group UK London. DOI: https://doi.org/10.1038/s41598-023-40743-z.

22. Nnaekwe, K., Ani, E., Obieke, V., Okechukwu, C., Usman, A., and Othman, M. (2025). Forecasting seasonal rainfall with time series, machine learning and deep learning. *Innovation in Computer and Data Sciences*, 1(1), 51–65. DOI: https://doi.org/10.64389/icds.2025.01127.

23. Sibanda, W. and Pretorius, P. (2012). Artificial neural networks-a review of applications of neural networks in the modeling of HIV epidemic. *International Journal of Computer Applications*, 44(16), 1–9. Publisher: Citeseer.

24. Asogwa, E. C., Nwankwo, M. P., Oguadimma, E. E., Okechukwu, C. P., and Suleiman, A. A. (2025). Hybrid LSTM-CNN deep learning framework for stock price prediction with google stock and reddit sentiment data. *Innovation in Computer and Data Sciences*, 1(1), 32–50. DOI: https://doi.org/10.64389/icds.2025.01126.

25. Bisaso, K. R., Karungi, S. A., Kiragga, A., Mukonzo, J. K., and Castelnuovo, B. (2018). A comparative study of logistic regression based machine learning techniques for prediction of early virological suppression in antiretroviral initiating HIV patients. *BMC medical informatics and decision making*, 18, 1–10. Publisher: Springer. DOI: https://doi.org/10.1186/s12911-018-0659-x.

26. Onyekwere, C. K., Nwankwo, C. K., Obulezi, O. J., and Ezeilo, C. I. (2025). The feature-value paradox: Unsupervised discovery of strategic archetypes in the smartphone market using machine learning. *Journal of Artificial Intelligence in Engineering Practice*, 2(2), 65–72. Publisher: The Scientific Association for Studies and Applied Research (SASAR). DOI: https://doi.org/10.21608/jaiep.2025.420689.1024.

27. Onyekwere, C. K., Nwankwo, C. K., Abonongo, J., Asogwa, E. C., and Shafiq, A. (2025). Economic growth dynamics: a machine learning-augmented nonlinear autoregressive distributed lag model of asymmetric effect. *Innovation in Computer and Data Sciences*, 1(1), 19–31. DOI: https://doi.org/10.64389/icds.2025.01125.

28. Shapley, L. S., *et al.* (1953). *A value for n-person games*. Publisher: Princeton University Press Princeton.

29. Haruna, U. D. and Samuel, M. (2024). An artificial neural network (ANN) model for predicting HIV infection among adolescents in Adamawa State. *Savannah Journal of Science and Engineering Technology*, 2(3), 103–109. Available at: https://www.sajsetjournal.com.ng/index.php/journal/article/view/102.

30. Seboka, B. T., Yehualashet, D. E., and Tesfa, G. A. (2023). Artificial intelligence and machine learning based prediction of viral load and CD4 status of people living with HIV (PLWH) on anti-retroviral treatment in Gedeo Zone Public Hospitals. *International Journal of General Medicine*. Pages: 435–451. Publisher: Taylor & Francis. DOI: https://doi.org/10.2147/IJGM.S397031.

31. Ugbor, G., Jamal, F., Khan, S., and Shawki, A. W. (2025). Generative AI for drug discovery: Accelerating molecular design with deep learning using Nigerian local content. *Innovation in Computer and Data Sciences*, 1(1), 66–77. DOI: https://doi.org/10.64389/icds.2025.01128.

Okechukwu J. Obulezi,
Department of Statistics,
Faculty of Physical Sciences,
Nnamdi Azikiwe University, P.O. Box 5025, Awka
Nigeria.
E-mail address: oj.obulezi@unizik.edu.ng

and

Mohamed A. F. Elbarkawy,
Department of Insurance and Risk Management,
College of Business, Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh, 11432,
Saudi Arabia.
E-mail address: maalbarqawi@imamu.edu.sa

and

Mmesoma P. Nwankwo,
Department of Statistics,
Faculty of Physical Sciences,
Nnamdi Azikiwe University, P.O. Box 5025 Awka,
Nigeria.
E-mail address: mp.nwankwno@stu.unizik.edu.ng

and

EbereChukwu Q. Chinedu,
Department of Statistics,
Faculty of Physical Sciences,
Nnamdi Azikiwe University, P.O. Box 5025 Awka,
Nigeria.
E-mail address: eq.chinedu@stu.unizik.edu.ng

and

Chinyere P. Igbokwe,
Department of Statistics,
School of Computer Science and Engineering, Lovely Professional University, Phagwara, Punjab, 144411, India
E-mail address: chinyere.12404612@lpu.in

and

Heba E. Salem,
Department of Management Information Systems,
College of Business and Economics, Qassim University,
Buraydah 51452, Saudi Arabia
Department of Mathematics,
Faculty of Science, Benha University,
Benha, Egypt.
E-mail address: H.Salem@qu.edu.sa

and

Mustafa Bayram,
Department of Computer Engineering,
Biruni Univetsity, 34010, Istanbul,
Turkey.
E-mail address: mustafabayram@biruni.edu.tr