RITAM ISH PUBLICATIONS

Journal of Progress in Artificial Intelligence

https://www.jpairitamish.org/

ISSN: 3139-4957 (Online)

Vol. 2, Iss. 4, October 2026 – December 2026

Successive Suppression of Viral Load in ART-Treated Patients: Machine Learning Algorithms and Dimensionality Techniques Prediction

Kabiru Bala[1], Ilker Etikan[2], Isah Muazu1[,3], Ibrahim Danladi Yakoko[4]

[1]Department of Statistics, School of Basic and Applied Sciences, Taraba State Polytechnic, P.M.B. 1030, Jalingo kabirubala7@gmail.com

[2]Near East University Faculty of Medicine, Head of Biostatistics Department, Near East Avenue, Postcode: 99138, Nicosia, North Cyprus, Mersin 10, Turkey ilker.etikan@neu.edu.tr

[3]Near East University Faculty of Medicine, Biostatistics Department, Near East Avenue, Postcode: 99138, Nicosia, North Cyprus, Mersin 10, Turkey isahisa2@gmail.com

[3]Department of Mathematics, Sokoto State University isahisa2@gmail.com

[4]Department of Curriculum and Instruction, College of Education, Zing, Taraba State, Nigeria diyakoko@gmail.com

Received: 23-Sep-2026, Manuscript No. JPAI-2026-0023; **Editor assigned:** 28-Sep-2026, PreQC No. JPAI-2026-0023 (PQ); **Reviewed:** 14-Oct-2026, QC No. JPAI-2026-0023; **Revised:** 19-Oct-2026, Manuscript No. JPAI-2026-0023 (R); **Published:** 12-Nov-2026

Citation: Bala, K., Etikan, I., Muazu, I., & Yakoko, I. D. (2026). Successive suppression of viral load in ART-treated patients: Machine learning algorithms and dimensionality techniques prediction. J Prog Artif Intell.

Corresponding: Kabiru Bala, Email: kabirubala7@gmail.com

OPEN ACCESS | RESEARCH ARTICLE

ABSTRACT

The use of Antiretroviral (ART) drugs significantly reduces Viral Load (VL) when adhered to properly. However, non-compliance can result in progression to Acquired Immune Deficiency Syndrome (AIDS) and potentially death. In this study, a combination of Machine Learning (ML) algorithms, Traditional Regression (TR), and Exploratory Data Analysis (EDA) was utilized for data modeling. The ML models included Multilayer Perceptron (MLP), Support Vector Machine (SVM), and Adaptive Neuro-Fuzzy Inference System (ANFIS). Traditional regression models consisted of Multiple Linear Regression (MLR) for predictive purposes and Logistic Regression (LR) for classification tasks. Multiple Correspondence Analysis (MCA) of EDA was employed to predict distance, covering model accuracy. The study utilized simple random sampling to select 1500 HIV patients on ART at Federal Teaching Hospital, Gombe State, Nigeria (FTHG). The MLP model demonstrated superior predictive performance in the calibration phase with an R-square (R^2) of 0.920209 and Mean Squared Error (MSE) of 0.002150. In the verification phase, MLR achieved a higher R^2 of 0.913961 and a lower MSE of 0.002234 than other models. The LR regression classifier model explained 73.3% of the variance in predicting patients' commitment to ART medication. Significant predictors included age, viral load, gender, marital status, age group, and drug type ($p < 0.05$). However, treatment compliance and defaulters demonstrated a sensitivity of 85.64%, specificity of 93.47%, positive predictive value of 87.92%, negative predictive value of 91.97%, and overall accuracy of 90.53%. The ROC curve indicated that the viral load score does not significantly predict outcomes ($p > 0.05$). Similarly, the MCA indicated an inertia of 45.5%, suggesting that the dimensions significantly contribute to total variance, with an eigenvalue of 5.47, and Cronbach's alpha was 69.1%, reflecting acceptable internal consistency.

KEYWORDS: Antiretroviral therapy, AI models, Discrimination measures, Sensitivity, Viral load

1.INTRODUCTION

Human Immunodeficiency Virus (HIV) causes Acquired Immune Deficiency Syndrome (AIDS) and is linked to high prevalence and depression among individuals on Antiretroviral Therapy (ART) in West Africa. Although ART allows HIV to be managed as a chronic condition, challenges such as long-term adherence, drug resistance, and access disparities persist. Recent research is focusing on alternative tactics such as long-acting injectable treatments and immunotherapeutic approaches to enhance treatment success and adherence, to achieve viral remission [1-3]. Early in the 20th century, HIV, a lentivirus from the Retroviridae family, was transmitted to humans in West-Central Africa from non-human primates [4]. To predict HIV cases, treatments, and drugs, several scholars used an extensive variety of computational models, ranging from ML and classical algorithms like random forests, support vector machines, artificial neural networks, multilayer perceptrons, and logistic regression. Scholars utilized these factors to predict HIV management, diagnostic tests, immunological results, and therapy monitoring [5], Consequently, quality of life in ART [6], Old age in HIV [7], Duration of viral suppression [8], Enhancing early diagnosis and treatment approaches [9] and Risk factors and treatment [10], ART medication among cardiovascular disease patients with HIV [11], ART drug circumstances [12],

CD4+T cells examined [13] and ART progress in HIV-1 [14]. In addition, MLP and ANN predicted HIV drug progress [15], and fuzzy logic investigated HIV diagnostic tools and challenges [16].

In the same vein, scholars unified neural networks and multivariate analysis to predict risk factors in HIV-infected populations [17]. Similarly, SVM, MLP, Random Forest, Cox regression, and eXtreme Gradient Boosting have been employed to investigate HIV/AIDS occurrences [18]. Classifier techniques of ML employed Logistic regression to investigate transmission of HIV-positive status in pregnant women [19], HIV Risk [20], Drivers of HIV testing uptake among women [21], and viral load in HIV children [22]. Logistic and Bayesian frameworks investigated late presentation of HIV infection [23], and Logistic regression and decision trees studied older adults living with HIV and other diseases [24].

Based on the foregoing, several data reduction techniques were also utilized by different scholars to evaluate HIV/AIDS cases. Scholars exploited PCA to study viral load [25], Factors connected to ART medication [26], and HIV type and CD4 [27]. Conversely, MCA and RF investigated viral load [28], MCA and cluster analysis studied patients on ART follow-up [29], MCA, logistic regression, and latent class analysis investigated HIV-related situations [30] and MDS, CFA, and chi-square were used to study ART-treated characteristics [31]. In this study, a synthesis of ML (predictive and non-predictive) and dimensionality techniques of MCA was developed to predict and evaluate viral load behavior with respect to ART treatment compliance.

2. MATERIAL AND METHOD

2.1 Design and Study Setting

The study combines machine learning, traditional regression, and Multiple Correspondence Analysis (MCA). The research was carried out at the Federal Teaching Hospital, Gombe State, Nigeria (FTHG). The hospital serves as a major tertiary referral centre for providing comprehensive HIV/AIDS treatment, care, support, and services within the north-eastern region of Nigeria.

2.2 Study Population

This study used simple random sampling to analyse the target population of HIV/AIDS patients undergoing active ART follow-up at FTHG. A total of 1500 HIV/AIDS patients were utilized for the study.

2.3 Data Collection Method

The study used a retrospective cohort study of ART patients to acquire the data. The ART unit management the downloaded data from the hospital data base of the hospital. The variables retrieved from the database include. Demographic profile of patients (age, age group, sex, marital status), Clinical variables (patients' weight, viral load, systolic and diastolic blood pressures), and Treatment variables (ART drugs).

2.4 Operationalization of Models Based On Different Independent Variables

2.4.1 Machine learning predictive model

The predictive supervised machine learning algorithms of AI (MLR, SVM, ANFIS) used age in ART as the target variable, while the independent variables used were sex, marital status, age group, weight, viral load, ART drug, SBP, and DBP. Before validating the models, the dataset was portioned into calibration (70) and verification (30) phases

concurrently. (R^2) and MSE were used to measure the accuracy of the model.

2.4.2 Machine Learning Classifier Predictive Model

The classifier predictive supervised machine learning model of logistic regression utilized a binary dependent variable (compliance and non-compliance with ART treatment) as the outcome or target variable, while the independent variables used all the predictors, but only the significant predictors were presented (age, viral load, marital status, age group, and ART drug). The classification model was rated based on Nagelkerke R^2 and odds ratio, Sensitivity, Specificity, positive predictive value (PPV), negative predictive value (NPV), accuracy, and ROC-AUC.

2.4.3 Unsupervised Statistical Methods

Unsupervised statistical methods of MCA focus on discovering hidden patterns and relationships within a dataset. MCA is an EDA functioning without hypothesis assumptions between dependent and independent variables. The main aim is to reduce data dimensionality from high structure to low structure levels, employing inertia for variance analysis. It also calculates distances to uncover geometric associations in a lower-dimensional space, using metrics like Cronbach's alpha, eigenvalues, inertia, correlation matrices, and discrimination measures. However, the Cronbach's alpha is computed based on the dimension to measure reliability of internal consistency and to quantify how well the variables load on a single dimension and work together to measure a single underlying latent construct. The eigenvalue computes the amount of variance captured in designing the new geometrical (dimensional) axis. In addition, the eigenvalue computes the percentage of explained inertia as well as the axis assigned the eigenvalue. The inertia is symbolically referred to as the chi-square computation, which divides Burt's matrix by the total number of observations. It plays a role in quantifying how much available structural information is captured. The Discrimination measures use a correlation matrix of Pearson type to find associations between qualitative variables. The matrix or the associated discrimination measures are used to identify the strongest or weakest variables in any dimension formed.

2.5 Data Pre-Processing and Normalization

Before the actual modelling, the retrieved data were censored and pre-processed to handle the inconsistency and missing data. Then, normalization was executed to ascertain the efficiency and accuracy of the data. Categorical age data were transformed to form age groups using records into different variables. MATLAB. SPSS and Excel were the main software used for preprocessing and data analysis. MATLAB handled ML model validation; SPSS was used for classification prediction and multiple correspondence analysis. Excel aided their graphical representation. Figure 1 is the study methodology flowchart.

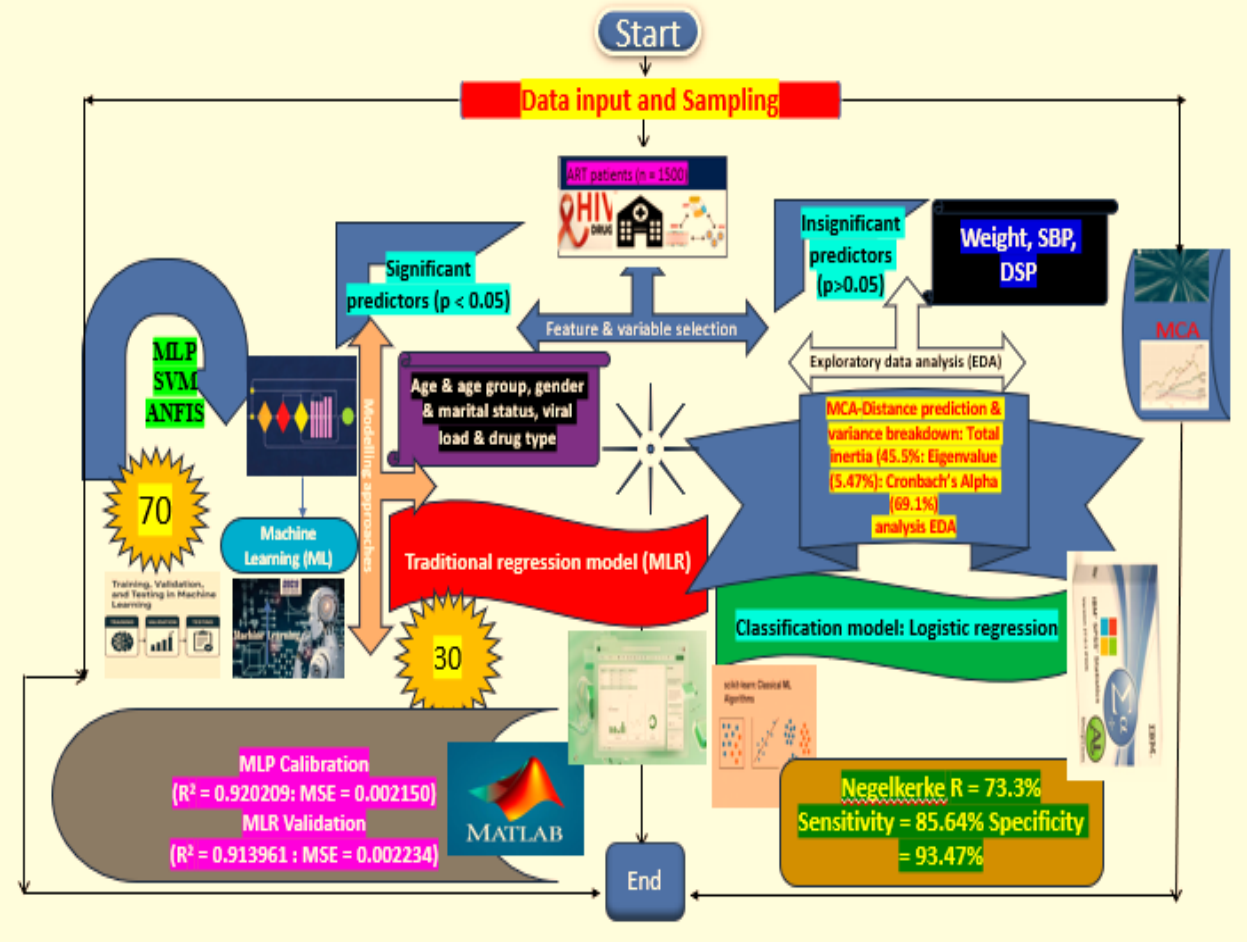


Figure 1. Architectural design and model evaluation workflow

3. Machine Learning Models

3.1 Multi-Layer Perceptron (MLP)

The network considered three concurrent layers: input, hidden, and output. Data enters at the input layer. Each node in the hidden layer connects to all nodes in the input layer and then to the output layer. This setup enables hidden neurons to facilitate computation between the input and output, with the output being the final layer of neurons [32]. The Multilayer Perceptron Neural Network (MLPNN) is defined as a structured network that processes and evaluates computational aspects through interconnected layers of composed neurons, managing the distribution of elements across these layers [33,34]. The MLPNN is specifically designed to address and model complex problems through interconnected nodes, making it a valuable tool for research in modeling and prediction. It has a unique technique of executing two different training patterns, which could be either forward or backward propagation algorithms [35]. Figure 2 is the graphical description of the MLP algorithm.

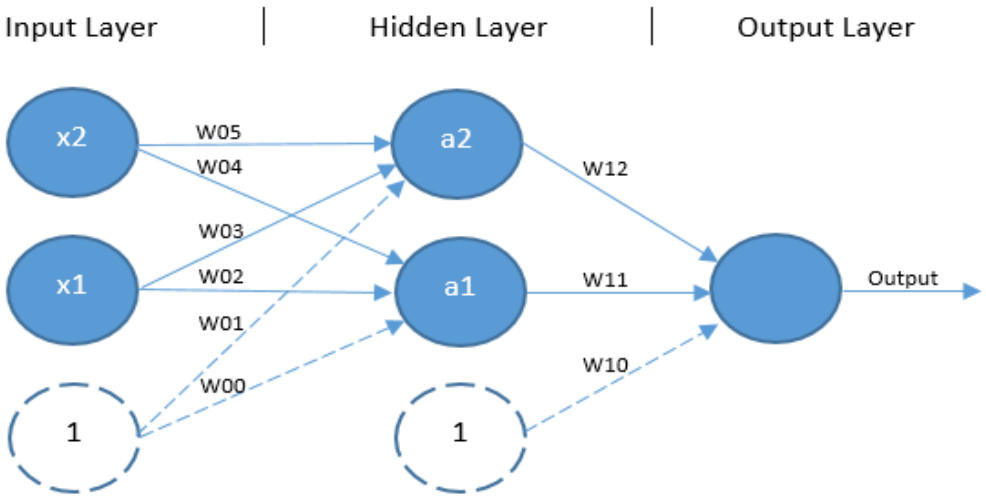


Figure 2. MLP design structure

3.2 Support Vector Machine (SVM)

The SVM is a classification algorithm designed to handle nonlinear data separation using training and testing datasets, intending to create a predictive model based on the laid procedure [36]. The uniqueness of the SVM model is empirically designed to minimize risk, resist overfitting, give high-performance results, and solve optimization problems [37]. The training sets are recruited as

$$(x_i\ ,\ y_i), i = l\ where\ x_i\ \epsilon\ R^m\ \ and\ y\ \epsilon\ \{1, -1\}^l \quad (1)$$

$$\min_{w,b,\xi}\ \frac{1}{2}w^T w + C\sum_{i=1}^{l}\xi_i \quad (2)$$

Subject to $y_i(w^T\phi\ (x_i) + b\ )\ \geq 1 - \xi_i$ (3)

$\xi_i\ \geq\ 0.$

Where;

x_i Signify the training vector that can handle the mapping to a higher-dimensional space through the function ϕ. Then the linear separating hyperplane alongside the maximal margin in the higher-dimensional space is being discovered by the SVM. $C > 0$, Symbolize the penalty parameter of the error [38]. Figure 3 is the SVM illustration design.

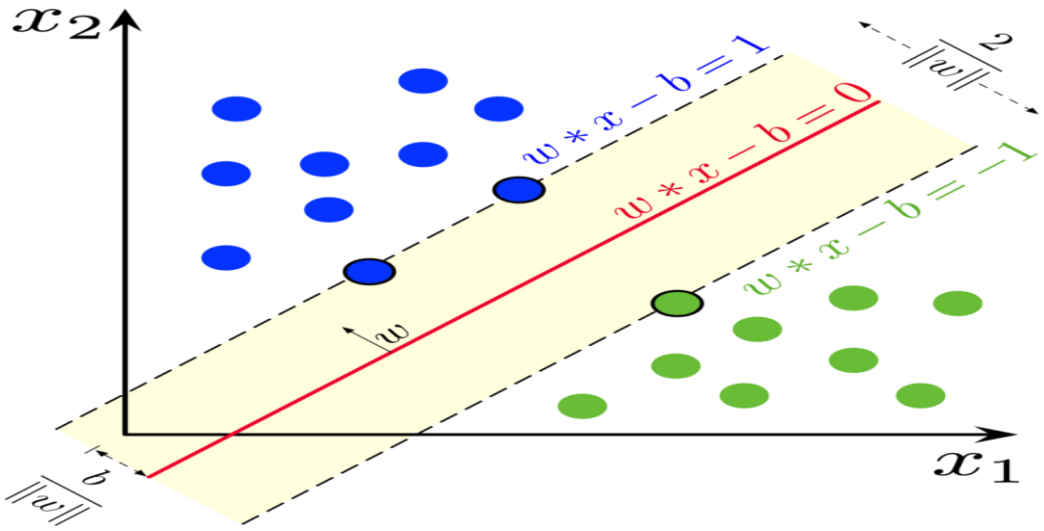


Figure 3. Hyperplane of support vector

3.3 Neuro-Fuzzy (NF)

The neuro-fuzzy technique serves as a powerful system that understands and interprets behavior very easily and possesses good learning capabilities. It lacks interpretation skills when it comes to neural networks. In addition, it is specifically designed to present the best possible parameters within a minimum time interval [39,40]. Neuro-fuzzy utilizes artificial intelligence to help data parameters learn and create algorithms for neural networks, addressing uncertainty and inaccuracies in judgments produced by these networks [41]. (See Figure 4).

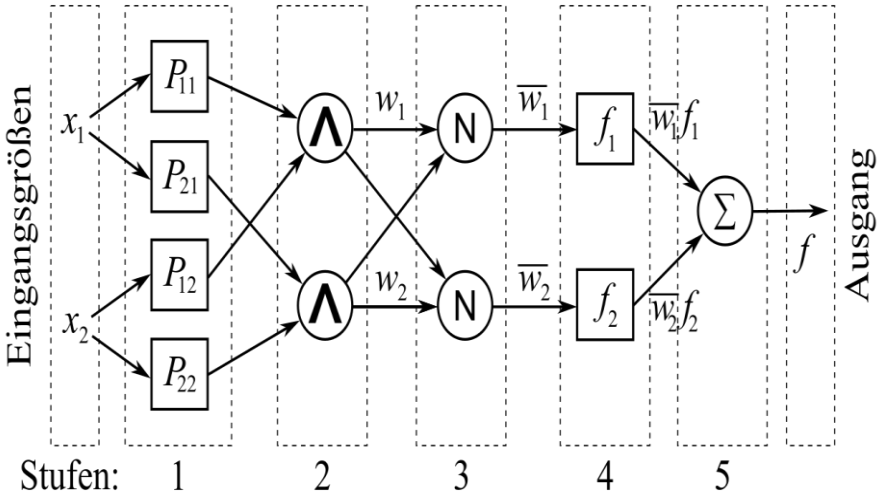


Figure 4. The ANFIS structural design

3.4 Multilinear Regression (MLR)

MLR utilizes a Gaussian distribution for residuals, centered on a mean variable that is linearly dependent on a specific set of coefficients. It serves as a dual function of linear modelling, addressing the relationship between a dependent variable and multiple independent variables [42]. The MLR sketching is described in Figure 5.

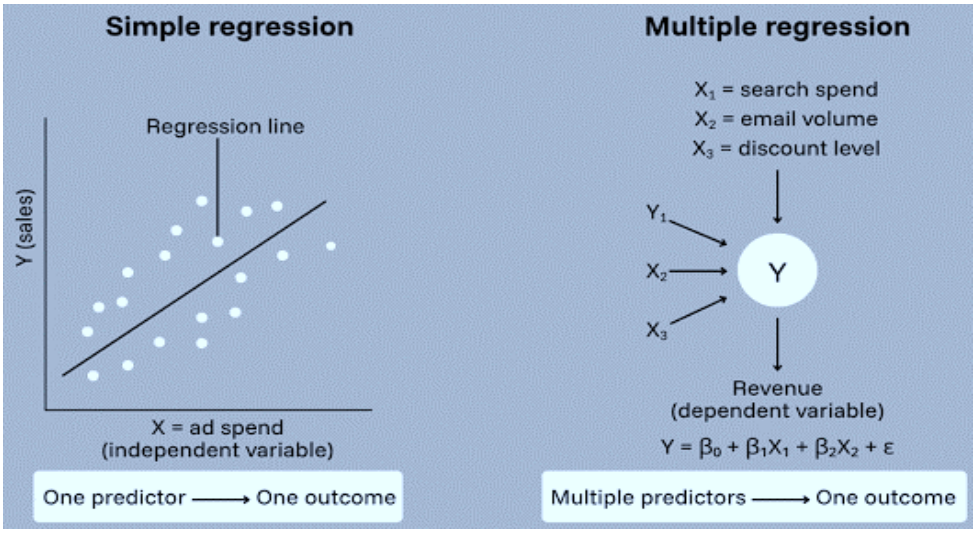


Figure 5. MLR sketch

3.5 Logistic Regression

The background of logistic regression (Bernoulli probability distribution) is within the capacity of a binary response variable Y (0 or 1), a linear combination of independent variables (Xs), and a logistic link function (complementary log-log, probit, logit). The log-odds are determined by the logit link function [43,44]. The sigmoid and linear diagram for the logistic regression is displayed in Figure 6.

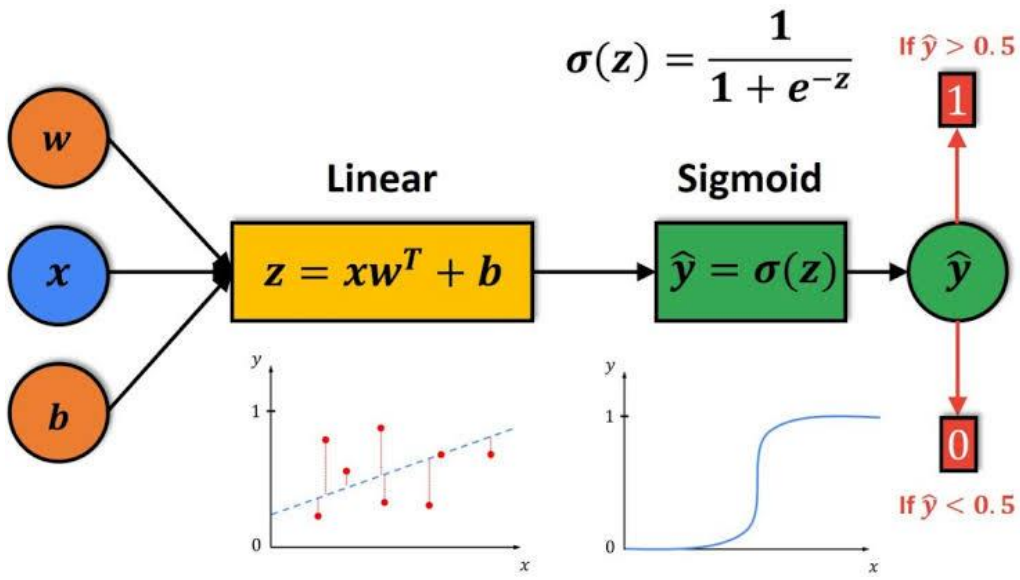


Figure 6. Outline of logistic regression

4. RESULTS AND DISCUSSION OF ML ALGORITHMS AND MCA MODELS

4.1 Model Performance of MLP, SVM, ANFIS, and MLR

In this section, ML algorithms utilized AI models to predict the number of years in ART follow-up. The models' goodness of fit and variance in Table 1 show that MLP had the best performance in capturing a high coefficient of determination (DC) with a smaller MSE in the calibration phase (DC = 0.920209; MSE = 0.002150) than SVM, ANFIS, and MLR, while in the verification phase MLR (DC = 0.913961; MSE = 0.002234) outperformed MLP, SVM, and ANFIS, respectively. Scholars have reported high accuracy and performance in HIV and ART medication-related case discovery in different studies, including MLP predicting lower accuracy between 66% and 95% than other ML models [45]. ANN predicted 95% outstanding performance in ART drug modelling [46]. ANN (75% and 76%) in training and testing outperformed MLR in protease activity prediction for HIV [47]. A similar study found AI did best in predicting the prevention and intervention of HIV [48]. Figures 7 and 8 showcase the graphical view of the model performances.

Table 1. Training and Testing Goodness-Of-Fit and Variance Model Performance

	Calibration				Verification			
Model	**R^2**	**R**	**MSE**	**RMSE**	**R^2**	**R**	**MSE**	**RMSE**
MLP	0.920209	0.959275	0.002150	0.046375	0.912092	0.955035	0.002283	0.044778

SVM	0.919605	0.958960	0.002166	0.046550	0.913696	0.955874	0.002241	0.047348
ANFIS	0.919024	0.958657	0.002182	0.046718	0.913284	0.955659	0.002253	0.047461
MLR	0.917345	0.957781	0.002228	0.047201	0.913961	0.956013	0.002234	0.047275

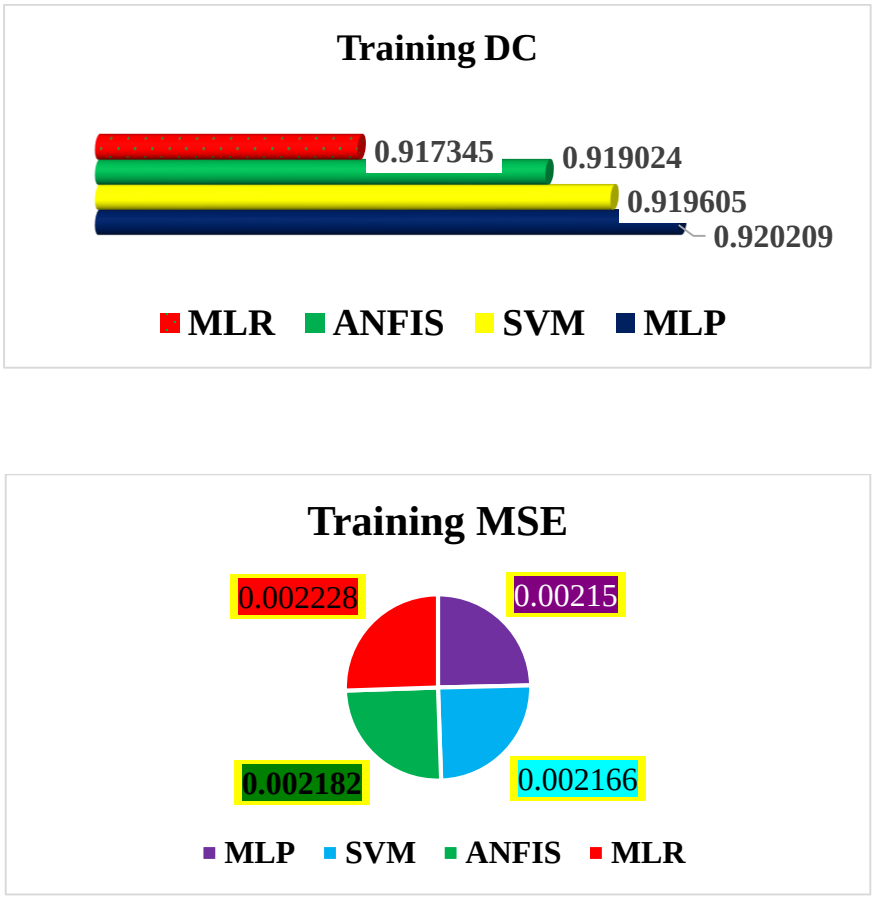


Figure 7. DC-MSE Training performance

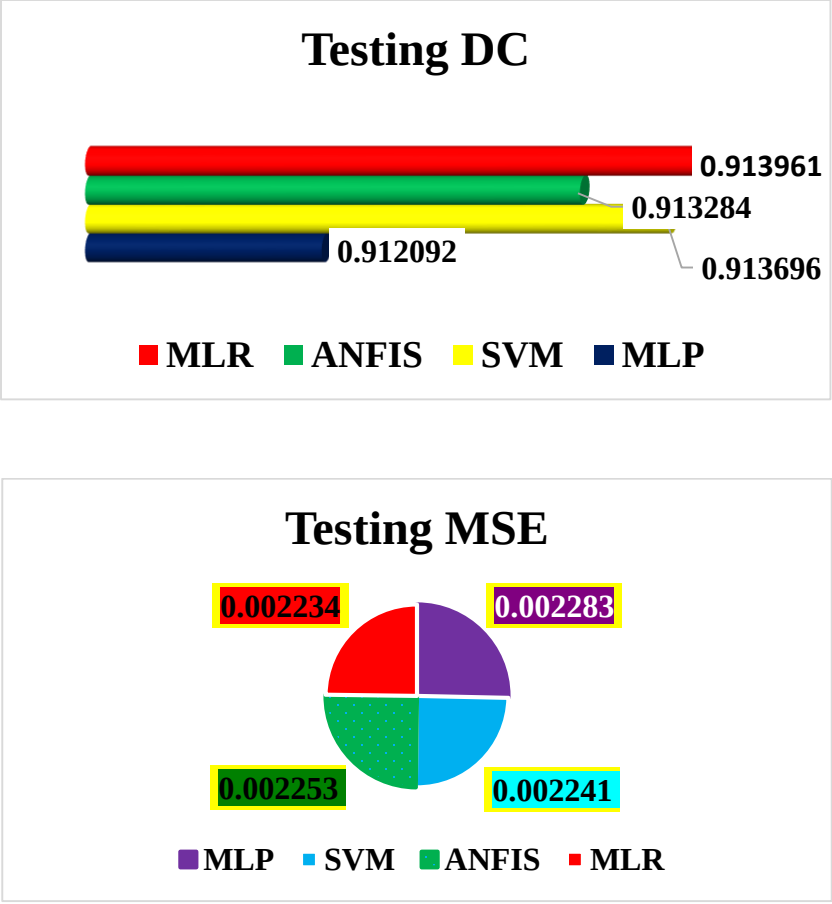


Figure 8. DC-MSE Testing Performance

4.2 Clinical Implications and Future Translational Framework:

Despite the strong performance exhibited by the proposed ML and traditional regression model, it serves as conceptual evidence of the framework. It cannot be fully implemented to assist in HIV//ART unless it undergoes multi-site testing and clinical validation. Effective validation can allow for incorporating the models into electronic health records (EHR) to serve as a policy formulation for clinical decision-making. During ART follow-up, healthcare professionals can use patients' bio data and clinical data to generate a real-time risk assessment for ART defaulters and compliance. This gives health professionals room to design allocations, interventions, counselling, and preventive risk to tackle the danger of viral load increase.

4.3 Model Performance of Logistic Regression

This section utilized logistic regression to predict classification accuracy in ART-treated patients. In Table 2, logistic regression was used to assess the effect of age, viral load, gender, marital status, age group, drug type, hospital status, SBP, DBP, and weight on the likelihood that patients will commit to or comply with ART treatment. The model explained 73.3% (Nagelkerke R^2) of the variance in compliance with ART medication. However, commitment to medication during ART treatment varied by predictor.

The predictors age (0.001), viral load (0.003), gender (0.001), marital status (0.001), age group (0.001), and drug type (0.025) added significantly to the model ($p < 0.05$). However, hospital status, SBP, DBP, and weight did not add significantly to the model ($p > 0.05$). The findings, based on odds, demonstrated that the prevalence was lower in married couples (OR = 0.883; 95%CI = 0.883 – 0.627; p = 0.001) compared to divorced and widowed. Previous studies found that the HIV prevalence was lower in the marital status of homosexual individuals (Qin et al., 2021; Su et al., 2016). However, the prevalence in age groups was reported higher in age group 23-32 (OR = 0.459; 95%CI = 0.099 – 2.119; p = 0.318), followed by 43-52 (OR = 0.298; 95CI = 0.127 – 0.698; p =0.05), 33-42 (OR = 0.270; 95%CI 0.084 – 0.866; p = 0.028) and 13-22 (OR = 0.250; 95%CI = 0.035 – 1.778; p = 0.166). A similar study also found that the prevalence of HIV was higher in older age than in younger age [51]. Adherence to ART treatment found age, gender, and family income were not associated [52]. Supervised machine learning found Age, regimen, WHO clinical stage, nutritional status, address status, sex, weight, recent CD4 cell count, viral load, and ART were the most important predictors for adherence status [53]. In HIV treatment, ART duration, BMI, and age were found to be positively associated with treatment; however, the use of condoms and functional status had negative effects on treatment [54,57].

Table 2. Significant predictors of logistic regression model performance

Parameters	**Exp (B)**	**95%CI (Lower-upper)**	**p-value**
Age	0.908	0.866 – 0.953	0.001
Viral load	1.000	1.000 -1.000	0.003
Marital status (Single)			0.001
Divorced	15.800	3.530 – 70.718	0.001
Married	0.883	0.883 – 0.627	0.001
Widowed	28.927	11.582 -72.249	0.478
Age group (53+)			0.001
13 – 22	0.250	0.035 – 1.778	0.166
23 – 32	0.459	0.099 – 2.119	0.318
33 – 42	0.270	0.084 – 0.866	0.028
43 – 52	0.298	0.127 – 0.698	0.005
ART drug (TDF-FTC-NPV)			0.025
ABC-DD1-LPV/r;	1.138	0.663 -1.954	0.639
AZT/3TC-ATV/r	1.133	0.548 – 2.342	0.736

AZT/3TC/EFV	1.599	0.772 – 3.311	0.206
AZT-3TC-LPV/r	2.159	1.020 – 4.572	0.044
AZT/-3TC-NPV	1.962	0.995 – 3.871	0.052
TDF-3TC-ATV/r	1.200	0.611 – 2.359	0.596
TDF-3TC-EFV	0.949	0.483 – 1.865	0.880
TDF-3TC-LPV/r	1.435	0.705 – 2.918	0.319
TDF-3TC-NPV	0.652	0.322 – 1.321	0.235
TDF-FTC-ATV/r	0.670	0.350 – 1.284	0.227
TDF-FTC-EFV	1.165	0.588 – 2.308	0.662
TDF-FTC-LPV/r	1.355	0.684 – 2.685	0.383

4.3.1 Validity diagnostic test

In this section, sensitivity (True positive-TP), specificity (True negative-TN), False Positives (FP), false negatives (FN), Positive Predictive Value (PPV), Negative Predictive Value (NPPV) and accuracy were utilized to evaluate compliance and non-compliance with antiretroviral therapy medication. In Table 3, the results demonstrated a sensitivity of 85.64% that patients on ART were truly complying with medication, while a specificity of 93.47% showed that patients were not complying with treatment. In addition, 6.8% were classified as false positives and 14.4% were classified as false negatives. The study demonstrated a positive predictive value of 87.92% for complying with treatment and a negative predictive value of 91.97%%, for non-compliance.

The Receiver Operating Characteristic Curve (ROC) coverage was 0.519 Area Under the Curve (AUC). Disclosing the viral load score cannot substantially predict the outcome variable, as indicated by the significance p-value of 0.214. The ROC in Figure 9 and the confusion matrix in Figure 10 further illustrate the area and how the variable is correlated.

Table 3. Classification of treatment status

	Compliance	**Defaulters**	**Total**
Compliance	465 True positives	64 False positives	529
Defaulters	78 False negatives	893 True negatives	971
Total	543	957	1500

Sensitivity (True positive) = TP/(TP+FN) = 465(465+78) = 465/543 = 85.64%

Specificity (True negative) = TN(TN+FP) = 893(893+64) = 893/957 = 93.47%

Positive predictive value (PPV) = TP(TP+FP) = 465/(465+64) = 465/529 = 87.92%

Negative predictive value (NPP) = TN(TN+FN) = 893/(893+78) = 893/971 = 91.97%

Accuracy = TP+TN/(TP+TN+FP+FN) = 465+893/(465+893+64+78) = 1358/(1500) = 90.53%

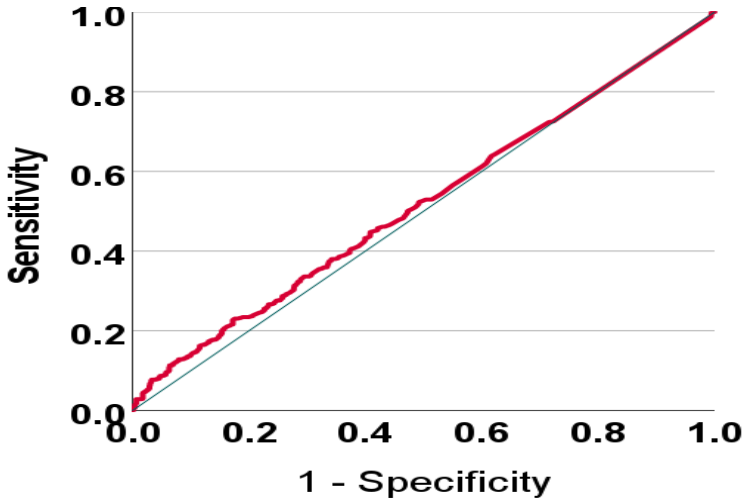


Figure 9. ROC illustration

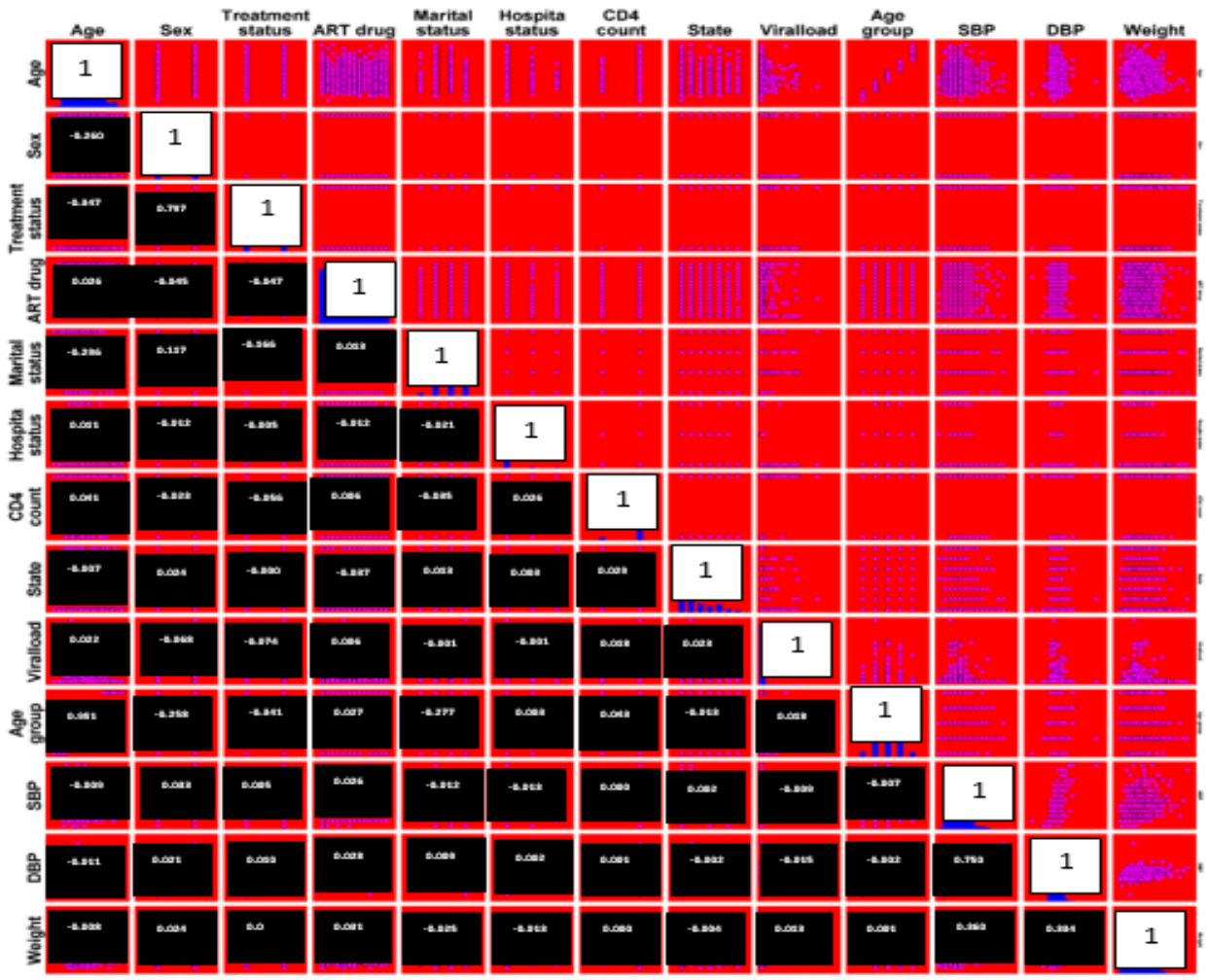


Figure 10. Variables confusion matrix

4.4 Model performance of multiple correspondence analysis

The unsupervised multivariate statistical method of MCA was used in this study to display the best geometrical axes and distances connected to ART contribution. The study results presented an acceptable internal consistency of 69.1% Cronbach's alpha and strong dimensional structure within the border limit. The eigenvalue explained 5.47 of the total inertia, indicating a meaningful, distinct pattern of associations among the categorical variable levels. However, the geometric axis design revealed that more information was captured in dimension I (2.848) compared to dimension II (2.613). The chi-square computed 45.5% inertia to determine the space captured by the new axis, indicating high structural heterogeneity information in dimension I (0.237) compared to dimension II (0.218).

The discrimination measures in this study mapped the geometrical axis in two dimensions. Dimension I indicate the strongest commitment to ART from weight (0.146), gender (0.196), marital status (0.269), viral load (0.478), age group (0.795), and age (0.822) while hospital status (0.001), State (0.004), CD4 count (0.005), ART drug (0.033), DSB (0.037), and SBP (0.062) from same dimension I indicate week

commitment to ART. Conversely, in dimension II the strongest compliance in ART was structural formed in DSB (0.275), SBP (0.294), weight (0.289), viral load (0.540), age group (0.564), and age (0.610), while week compliance formation was found in DSB (0.275), SBP (0.294), weight (0.289), viral load (0.540), age group (0.564), and age (0.610), respectively. The overall strongest variables associated with ART treatment compliance in both dimensions were found in age (0.716), age group (0.6795), and viral load (0.509). Scholars investigated using MCA and found that age group, ART drug, gender, marital status, and educational level were not associated with ART treatment adherence [55]. Figure 11 is the structural dimensional axis plot.

Table 5. Contributions of discrimination measures

Variables	**Dimension**		**Total and Mean**
	1	2	
Age	0.822	0.610	(1.432 - 0.716)
Gender	0.196	0.004	(0.2 - 0.100)
ART drugs	0.033	0.022	(0.055 - 0.0275)
Marital status	0.269	0.003	(0.272 - 0.136)
Hospital status	0.001	0.004	(0.005 - 0.0025)
CD4count	0.005	0.001	(0.006 - 0.003)
State	0.004	0.008	(0.012 - 0.006)
Age group	0.795	0.564	(1.359 - 0.6795)
Weight	0.146	0.289	(0.435 - 0.2175)
SBP	0.062	0.294	(0.356 - 0.178)
DBP	0.037	0.275	(0.312 - 0.156)
Viral load	0.478	0.540	(1.018 - 0.509)
Active Total	2.848	2.613	(5.412 - 2.731)
Cronbachs Alpha	0.708	0.673	(1.381 - 0.691)
Eigen value	2.848	2.613	(5.462 - 2.731)
Inertia	0.237	0.218	(0.455 - 0.228)

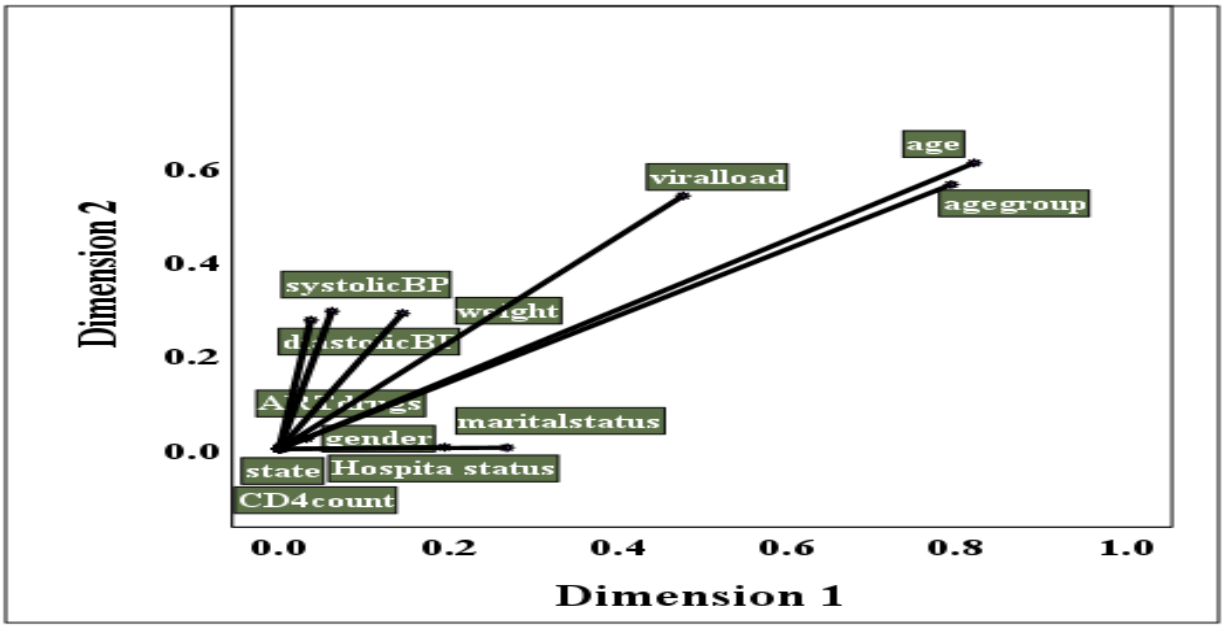


Figure 11. MCA dimensionality structure of HIV/ART patients

5. CONCLUSION

The successive suppression of viral load prediction and analysis utilizing HIV patients yielded significant results in this study. In training, MLP surpassed SVM, ANFIS, and MLR, but in testing, MLR superseded MLP, SVM, and ANFIS, respectively. ML algorithms are modernizing HIV research in terms of providing means to predict viral mutations and improve drug innovation [56]. The predictors age, viral load, gender, marital status, age group, and drug types made a significant contribution to the model ($p > 0.05$), allowing the logistic regression to explain 73.3% of the variance in the model. The model further revealed 85.64% sensitivity, 93.47% specificity, 87.92% PPV, 91.97% NPV and an accuracy of 90.53%. The ROC showed that, with treatment status is the same, viral load score cannot significantly predict the outcome variable ($p > 0.05$). The MCA yielded an acceptable internal consistency at a borderline level. Based on the foregoing, the utilization of MATLAB software in this study revealed the qualities of AI (XAI) analysis features in conducting permutation-based predictor analysis to allow the black-box nature to translate MLP as the highest-performing model. Thus, the ML models can use VL and other HIV factors to predict ART treatment compliance. However, the statistical regression yielded significant clinical importance in viral load and demographic variables. Likewise, insignificant importance values closer to zero were revealed for weight, SBP, and DBP. Thus, there is a need to validate these findings of logistic regression at p-values less than 0.05 and solidify the models to concentrate on clinical behaviours rather than just relying on physiological variances.

AUTHOR CONTRIBUTION: Study design: All authors; Data collection: K Bala and Isah Muazu; Data analysis and interpretation: All authors;

Manuscript preparation and review: All authors; Proofreading: Isah Muazu, I Etikan, Danladi I. Yakoko

COMPETING INTERESTS: The authors declare no conflict of interest.

ACKNOWLEDGMENT: The authors extend their regards to the ART Unit, Federal Teaching Hospital, Gombe State, for granting ethical approval and data availability to achieve this success

ETHICAL APPROVAL: Ethics and Federal Teaching Hospital, Gombe State, Nigeria

LIMITATIONS AND EXTERNAL VALIDATION OF THE STUDY

The limitation in this study is centered on a single internal dataset, retrieved from the ART unit of FTHG. The calibration and verification accuracy shown by MLP and MRL to demonstrate internal predictive performance does not guarantee the generalization of these findings, because external validation is required. Strong performance could be found using different hospitals within the zone due to varying factors like demographic data, patient population, ART compliance attitudes, motivation, facilities improvement, clinical logistics, and resource availability. To improve clinical transformation and model health. In the future, researchers should expand the use of cross-validation of ML and traditional regression structural design to form a central dataset across diverse geopolitical regions.

6. REFERENCES

[1] Charpentier C, Berzow D, Le Hingrat Q, Kaiser R, Gomes P, Miranda AC, et al. HIV-2 EU—supporting standardized HIV-2 drug resistance interpretation: An update. Clin Infect Dis. 2025;81(3):539-542. doi:10.1093/cid/ciaf110.

[2] Łupina K, Nowak K, Lorek D, Nowak A, Romac A, Głowacka E, et al. Pharmacological advances in HIV treatment: from ART to long-acting injectable therapies. Arch Virol. 2025;170(9). doi:10.1007/s00705-025-06381-8.

[3] Magomere E, Olwal CO, Tetteh BE, Appeaning M, Ndung'u T, Kyei GB, et al. The confluence of HIV-1 and HIV-2: implications for disease progression and insights for therapy. Int J Microbiol. 2025;2025(1). doi:10.1155/ijm/3145677.

[4] Omondi FH, Fredrick C, Omondi H. Leveraging within-host HIV diversity to understand HIV persistence and proviral evolutionary dynamics in understudied populations. 2026.

[5] Ngcobo S, Mntla EM, Shock J, Louw M, Mbonambi L, Serite T, et al. Artificial intelligence for HIV care: a global systematic review of current studies and emerging trends. J Int AIDS Soc. 2025;28:70045. doi:10.1002/jia2.70045/full.

[6] Xie M, Zhang Z, Yu Y, Zhang L, Zhang J, Wu D. Predicting quality of life in people living with HIV: a machine learning model integrating multidimensional determinants. Health Qual Life Outcomes. 2025;23(1). doi:10.1186/s12955-025-02398-4.

[7] Xu Y, Li M, Su C, Zhu Q, Liu Q, Zhang Y, et al. Associated factors of cognitive frailty in people living with HIV aged 50 and older: a cross-sectional study. Infect Dis Ther. 2025;14(10):2429-2444. doi:10.1007/s40121-025-01218-y.

[8] Orowe I. Use of machine learning models to predict viral suppression and associated factors.

Appl Math Sci. 2025;19(8):359-369. doi:10.12988/ams.2025.919280.

[9] Senel B, Beka H. Artificial intelligence in HIV diagnosis and treatment: a comprehensive review. Cumhuriyet Med J. 2025;47(1):10-21. doi:10.7197/cmj.1630164.

[10] Barznji NSK. Prediction by sigmoid multilayer perceptron artificial neural network function and model selection for the risk factors most affected human immunodeficiency. Zanco J Humanit Sci. 2025;29(4):272-293. doi:10.21271/zjhs.29.4.13.

[11] Cernuschi AL, Masci C, Corso F, Muccini C, Ceccarelli D, Galli L, et al. A neural-network approach for predicting time to cardiovascular diseases in HIV patients based on real-world data. Oper Res. 2025;25(4). doi:10.1007/s12351-025-00972-8.

[12] Oyediran KO, Cardoso-Daodu IM, Bassey PO, Ogundemuren DA, Muhammed R, Ojo OE, et al. Artificial neural network guided optimization of tenofovir and Cyanovirin-N multipurpose preventive hydrogel formulation developed using the Box Behnken design model. Pharm Sci Adv. 2025;3:100079. doi:10.1016/j.pscia.2025.100079.

[13] Riaz FM, Ahmad S, Khan JA, Altaf S, Rehman ZU, Memon SK. Numerical treatment of non-linear system for latently infected CD4+T cells: a swarm-optimized neural network approach. IEEE Access. 2024;12:103119-103132. doi:10.1109/ACCESS.2024.3429279.

[14] Di Teodoro G, Siciliano F, Guarrasi V, Vandamme AM, Ghisetti V, Sönnerborg A, et al. A graph neural network-based model with out-of-distribution robustness for enhancing antiretroviral therapy outcome prediction for HIV-1. Comput Med Imaging Graph. 2025;120:102484. doi:10.1016/j.compmedimag.2024.102484.

[15] Ali SM, Ghazvinian A, Feng B, Feng J. Enhancing efficiency in the healthcare sector through multi-objective optimization of freight cost and delivery time in the HIV drug supply chain using machine learning. Systems. 2025. doi:10.3390/systems.

[16] Oluwagbemi OO, Onyeka E, Jatto A, Oluwagbemi FE, Smith S, Oyelami O, et al. Multilingual voice-enabled informatics tools: catalyst for equitable AI in HIV and HIV-comorbidity healthcare management. PLoS One. 2025;20(10). doi:10.1371/journal.pone.0332573.

[17] Li X, Wang JJ, Quan XM, Zhao CS. A multivariate analysis of the impact of HIV infection on the risk of osteoporosis based on the NHANES database. Medicine (Baltimore). 2025;104(36):e44459. doi:10.1097/MD.0000000000044459.

[18] Cai Q, Yang L, Ling Y, Pan W, Zhong Q, Wang C, et al. Survival prediction models for people living with HIV based on four machine learning models. Sci Rep. 2025;15(1). doi:10.1038/s41598-025-16479-3.

[19] Minda Z, Oljira L. HIV vertical transmission linked to virological failure in pregnant women on ART. AIDS Res Ther. 2025;22(1). doi:10.1186/s12981-025-00820-x.

[20] Albernas A, Patel MD, Cook RL, Vaddiparti K, Prosperi M, Liu Y. HIV risk score and prediction model in the United States: A scoping review. AIDS Behav. 2025;29(8):2388-2407. doi:10.1007/s10461-025-04702-1.

[21] Nketiah-Amponsah E, Ampaw S, Bawuah A. Rural-urban determinants of HIV/AIDS testing uptake among Ghanaian women of reproductive

age. BMC Public Health. 2025;25(1). doi:10.1186/s12889-025-24072-6.

[22] Ndlovu B, Kiwa FJ, Muduva M, Chipfumbu CT, Marambi S. Developing a logistic regression machine learning model that predicts viral load outcomes for children living with HIV in Gutu District, Zimbabwe. Indones J Innov Appl Sci. 2025;5(3):277-304. doi:10.47540/ijias.v5i3.2275.

[23] Zhao HH, Luo DH, Fei LP, Wang S, Chen FF, Qin QQ, et al. Factors influencing late HIV presentation in China: results from logistic regression and Bayesian network analyses. BMC Infect Dis. 2026;26(1). doi:10.1186/s12879-025-12429-6.

[24] Ge C, Gao H, Xia L, Wang Y, Zou M, Qian X. Predictors of mild cognitive impairment in older adults living with HIV and multimorbidity: A comparative analysis using logistic regression and decision tree models. Front Public Health. 2026;14. doi:10.3389/fpubh.2026.1837495.

[25] Ma Y, Yang X, Xiao J, Li X, Olatosi B, Zhang J. Longitudinal viral load clustering for people with HIV using functional principal component analysis. AIDS Res Treat. 2025;2025(1). doi:10.1155/arat/5890464.

[26] Nikolaev AA. Epidemic engines: a study of the factors driving HIV infection spread and mortality in the regions of the Russian Federation. Popul Econ. 2025;9(1):75-96. doi:10.3897/popecon.9.e118182.

[27] Mishra SK, Kumar N, Tóth Z, Bin Jardan YA, Akash S, Georrge JJ. Unveiling the molecular activity of HIV towards the CD4: a study based on subtype C via docking and dynamics approach. J Genet Eng Biotechnol. 2025;23(1):100457. doi:10.1016/j.jgeb.2025.100457.

[28] Soogun AO, Kharsany ABM, Zewotir T, North D, Ogunsakin RE. Identifying potential factors associated with high HIV viral load in KwaZulu-Natal, South Africa using multiple correspondence analysis and random forest analysis. BMC Med Res Methodol. 2022;22(1). doi:10.1186/s12874-022-01625-6.

[29] da Silva Ziani J, da Silva Corcini LMC, Härter J, Alves CF, Scholze AR, Padoin SM de M, et al. Characteristics of people hospitalized for HIV: Multivariate cluster analysis of antiretroviral therapy interruption. Rev Gaucha Enferm. 2025;46(Spec Issue 1). doi:10.1590/1983-1447.2025.20250125.en.

[30] El-Osta A, Bagkeris E. Public attitudes toward pharmacy-based HIV PrEP in the UK: a national cross-sectional study of 15,000 NHS patients. 2026. doi:10.21203/rs.3.rs-8466343/v1.

[31] Kieu PT, Vuong NL, Dung DV. Validation of multidimensional scale of perceived social support (MSPSS) in Vietnamese among people living with HIV/AIDS. AIDS Behav. 2023;27(8):2488-2496. doi:10.1007/s10461-022-03974-1.

[32] Tien Bui D, Tuan TA, Klempe H, Pradhan B, Revhaug I. Spatial prediction models for shallow landslide hazards: a comparative assessment of the efficacy of support vector machines, artificial neural networks, kernel logistic regression, and logistic model tree. Landslides. 2016;13(2):361-378. doi:10.1007/s10346-015-0557-6.

[33] Kavzoglu T, Mather PM. The use of backpropagating artificial neural networks in land cover classification. Int J Remote Sens. 2003;24(23):4907-4938. doi:10.1080/0143116031000114851.

[34] Ojha VK, Abraham A, Snášel V. Metaheuristic design of feedforward neural networks: A review of two decades of research. Eng Appl Artif Intell. 2017. doi:10.1016/j.engappai.2017.01.013.

[35] Pham BT, Tien Bui D, Prakash I, Dholakia MB. Hybrid integration of multilayer perceptron neural networks and machine learning ensembles for landslide susceptibility assessment at Himalayan area (India) using GIS. Catena. 2017;149:52-63. doi:10.1016/j.catena.2016.09.007.

[36] Yang C, Odvody GN, Fernandez CJ, Landivar JA, Minzenmayer RR, Nichols RL. Evaluating unsupervised and supervised image classification methods for mapping cotton root rot. Precis Agric. 2015;16(2):201-215. doi:10.1007/s11119-014-9370-9.

[37] Du XF, Leung SCH, Zhang JL, Lai KK. Demand forecasting of perishable farm products using support vector machine. Int J Syst Sci. 2013;44(3):556-567. doi:10.1080/00207721.2011.617888.

[38] Mierswa I. Evolutionary learning with kernels. 2006. p.1553. doi:10.1145/1143997.1144249.

[39] Łapa K, Cpałka K, Rutkowski L. New aspects of interpretability of fuzzy systems for nonlinear modeling. Stud Comput Intell. 2018;738. doi:10.1007/978-3-319-67946-4_9.

[40] Gago J, Martínez-Núñez L, Landín M, Flexas J, Gallego PP. Modeling the effects of light and sucrose on in vitro propagated plants: A multiscale system analysis using artificial intelligence technology. PLoS One. 2014;9(1). doi:10.1371/journal.pone.0085989.

[41] Abba SI, Hadi SJ, Abdullahi J. River water modelling prediction using multi-linear regression, artificial neural network, and adaptive neuro-fuzzy inference system techniques. Procedia Comput Sci. 2017;120:75-82. doi:10.1016/j.procs.2017.11.212.

[42] Grégoire G. Multiple linear regression. EAS Publ Ser. 2015;66:45-72. doi:10.1051/eas/1466005.

[43] Mustafa I, Allam A. A computer implementation of logistic regression without high-level statistical libraries. 2026.

[44] Zaidi A, Al Luhayb ASM. Two statistical approaches to justify the use of the logistic function in binary logistic regression. Math Probl Eng. 2023;2023(1). doi:10.1155/2023/5525675.

[45] Hie B, Zhong ED, Berger B, Bryson B. Learning the language of viral evolution and escape. 2009. Available from: http://science.sciencemag.org/

[46] Steiner MC, Gibson KM, Crandall KA. Drug resistance prediction using deep learning techniques on HIV-1 sequence data. Viruses. 2020;12(5). doi:10.3390/v12050560.

[47] Deeb O, Jawabreh M. Exploring QSARs for inhibitory activity of cyclic urea and nonpeptide-cyclic cyanoguanidine derivatives HIV-1 protease inhibitors by artificial neural network. Adv Chem Eng Sci. 2012;2(1):82-100. doi:10.4236/aces.2012.21010.

[48] Xiang Y, Du J, Fujimoto K, Li F, Schneider J, Tao C. Application of artificial intelligence and machine learning for HIV prevention interventions. Lancet HIV. 2022;9(1):e54-e62. doi:10.1016/S2352-3018(21)00247-2.

[49] Qin S, Lai J, Zhang H, Wei D, Lv Q, Pan X, et al. Predictive factors of viral load high-risk events for virological failure in HIV/AIDS patients receiving long-term antiviral therapy. BMC Infect Dis. 2021;21(1). doi:10.1186/s12879-021-06162-z.

[50] Su S, Chen X, Mao L, He J, Wei X, Jing J, et al. Superior effects of antiretroviral treatment among

men who have sex with men compared to other HIV at-risk populations in a large cohort study in Hunan, China. Int J Environ Res Public Health. 2016;13(3). doi:10.3390/ijerph13030283.

[51] Chireshe E, Chifurira R, Chinhamu K, Batidzirai JM, Kharsany ABM. Spatial analysis of HIV determinants among females aged 15–34 in KwaZulu Natal, South Africa: A Bayesian spatial logistic regression model. Int J Environ Res Public Health. 2025;22(3). doi:10.3390/ijerph22030446.

[52] Pinheiro CAT, De-Carvalho-Leite JC, Drachler ML, Silveira VL. Factors associated with adherence to antiretroviral therapy in HIV/AIDS patients: a cross-sectional study in Southern Brazil. Braz J Med Biol Res. 2002;35(10).

[53] Yeneakal KA, Teferi GH, Mihret TT, Mengistu AK, Tizie SB, Tadele MM. Predicting antiretroviral therapy adherence status of adult HIV-positive patients using machine-learning Northwest, Ethiopia, 2025. BMC Med Inform Decis Mak. 2025;25(1). doi:10.1186/s12911-025-03106-4.

[54] Mengistu AK, Yeneakale KA, Baykemagn ND, Melese ZY, Gedefaw AE. Application of causal forest double machine learning (DML) approach to assess tuberculosis preventive therapy's impact on ART adherence. Sci Rep. 2025;15(1). doi:10.1038/s41598-025-14460-8.

[55] Dieckhaus KD, Odesina V. Outcomes of a multifaceted medication adherence intervention for HIV-positive patients. AIDS Patient Care STDS. 2007;21(2):81-91. doi:10.1089/apc.2006.0044.

[56] Oyediran KO, Bassey POO, Ogundemuren DA, Abdulraheem A, Azubuike CP, Amenaghawon AN, et al. Artificial intelligence in human immunodeficiency virus mutation prediction and drug design: advancing personalized treatment and prevention. Pharm Sci Adv. 2025;3:100080. doi:10.1016/j.pscia.2025.100080.

[57] Quang Tu N. Characteristics of new mutations in the integrase gene and in the 3'PPT of the human immunodeficiency virus in the resistance mechanism [thesis]. Available from: https://theses.hal.science/tel-05612916v1