Prediction of CD4 T-Lymphocyte count via Machine Learning for HIV-positive patients

Saad Lamjadli[1,*], Oumayma Ouedrhiri[2], Ikram Souli[1], Zouhair Elamrani Abou Elassad[4], Oumayma Banouar[2], Safa Machraoui[1], Moulay Yassine Belghali[1], Raja Hazime[1], Noura Tassi[3], Said Raghay[2], Brahim Admou[1,5]

[1] Laboratory of Immunology, Center of Clinical Research, Mohammed VI University Hospital, Marrakech, Morocco
[2] Laboratory of Computer Science Engineering. Marrakech, Morocco
[3] Infectious Diseases Department, CHU Mohamed VI. Marrakech, Morocco
[4] SARS Research Team, Computer Science Department, ENSAS, Cadi Ayyad University, Marrakech, Morocco
[5] Biosciences Research Laboratory, Faculty of Medicine and Pharmacy, Cadi Ayyad University, Marrakech, Morocco
Email: saad.lamjadli@edu.uca.ma (S.L); oumayma.ouedrhiri@ced.uca.ma (O.O); ikramsouli90@gmail.com (I.S); z.elamrani@uca.ma (Z.E.A.E); o.banouar@edu.uca.ac.ma (O.B); s.machraoui.ced@uca.ac.ma (S.M); belghali@hotmail.com (M.Y.B); raja.hazime@hotmail.com (R.H); n.tassi@yahoo.fr (N.T); s.raghay@uca.ac.ma (S.R); br.admou@uca.ac.ma (B.A)
[*]Corresponding author

Manuscript received May 12, 2024; revised June 24, 2024; accepted July 4, 2024; published September 25, 2024

***Abstract*—The World Health Organization recommends routine immunological and virologic monitoring for all patients with Human Immunodeficiency Virus (HIV) infection. However, viral load and lymphocyte T CD4 (LTCD4) count analysis requires sophisticated equipment and qualified human resources. This creates a financial burden, especially in limited resource settings. Thus, there is a need for alternative approaches. One such alternative is machine learning (ML), which offers a more cost-effective solution. In this study, five highly optimized data-driven models for LTCD4 prediction were designed based on popular ML techniques: support vector machine (SVM), random forest (RF), logistic regression (LR), artificial neural networks (ANNs), and naive Bayes (NB). To guarantee the robust performance of the proposed algorithms, we meticulously scrutinized the optimal approach for constructing models. Furthermore, we analyzed the predictive capabilities of LTCD4 according to multiple thresholds of the total lymphocyte count. Moreover, an imbalance-aware strategy to overcome the aforementioned issue was adopted using the synthetic minority oversampling technique. The cutoff points for the number of lymphocytes "1100" had the best performance in predicting the LTCD4 count. SVMs, RF, NB, LR, and ANNs provided an area under the curves of 97%, 93.2%, 90%, 92.01%, and 93%, respectively. SVMs achieved better results in predicting LTCD4 in all metrics. The results offer novel perspectives on LTCD4 forecasting, presenting opportunities to enhance initiatives aimed at developing web-based systems. These systems could alleviate the financial burden associated with measuring LTCD4 in patients with HIV infection, particularly in resource-constrained settings.**

***Keywords*—HIV, CD4 T lymphocyte count, total lymphocyte count, machine learning**

I. Introduction

Human Immunodeficiency Virus is known to be one of the major concerns and threatening problems encountered by societies, resulting in many health issues, economic losses, and fatalities. The World Health Organization (WHO, 2023) reported that 40.4 million (32.9–51.3 million) people are currently living with ongoing HIV transmission worldwide [1]. Throughout HIV-related care and support programs, the LTCD4 counts are used for progression monitoring and treatment response assessment [2, 3]. Therefore, understanding the conditions under which changes in CD4 count significantly influence the development of effective policy interventions is crucial for minimizing variations in LTCD4 levels. Reliable LTCD4 count forecasting and proactive analysis are undeniably of great interest and necessity. Most studies have been conducted on LTCD4 variation. The latter is a complex mechanism influenced by multiple contributing factors such as the HIV viral load, antiretroviral therapy, coinfections, age, and adherence to medical care factors [4]. Although previous research offers valuable perspectives on fostering positive safety practices, it is imperative to acknowledge that during the LTCD4 count prediction process, diverse modeling techniques yield distinct performance measures. In this context, machine learning (ML) models have demonstrated superiority over statistical analyses in forecasting future events and have documented satisfactory outcomes [5]. Support vector machines (SVMs) and artificial neural networks (ANNs) are among the most substantial ML techniques that have been used for LTCD4 count prediction [6]. It has been asserted that the SVM model has the capability to manage small data sizes, exhibits excellent performance in mitigating overfitting issues, and demonstrates superior generalization abilities [7]. Conversely, ANNs have garnered widespread acclaim because of their proficiency in handling diverse and complex tasks. They have progressively gained recognition for their capacity to acquire data representations in both supervised and unsupervised settings, coupled with parallel processing, fault tolerance, and efficiency in generalizing to unseen data samples through hierarchical representations [8]. In the literature pertaining to LTCD4 prediction, techniques such as SVM and ANNs are recognized as prominent and effective because of their robust theoretical foundations. Nevertheless, these ML algorithms encounter a significant challenge in achieving optimal performance results, prompting the exploration of parameter optimization strategies. Previous research has demonstrated that for enhanced performance metrics in the SVM model, optimization of the penalty factor (C) and kernel parameter (c) is crucial. The optimization of SVM was executed using the grid search method, which is a widely adopted and proven efficient approach for tuning the model's hyperparameters [9]. Similarly, ANNs optimization can be realized by tuning the number of layers, input and hidden neurons, weights, etc. A trial-and-error strategy along with the dropout regularization

method [10] and early-stop approaches were used to achieve this objective. The k-fold cross-validation technique was adopted to evaluate the classification performance. It has been recognized for its susceptibility to yield minimal bias and variance in contrast to the other validation methods, including the leave-one-out method [11]. In the context of current research, SVM and ANNs model applications represent a significant advancement in statistical modeling. These techniques not only address the limitations of traditional methods but also align with the overarching objectives of statistical analysis to provide accurate, reliable, and interpretable results [6, 12]. By leveraging these advanced ML approaches, researchers can achieve more precise predictions and gain deeper insights into their data, thereby pushing the boundaries of what can be achieved through statistical modeling hence advancing our understanding and ability to predict future events [12]. Data necessary for LTCD4 count prediction systems can be gathered through various experimental formats, including numerical registries in hospitals [13]. Selecting data mining studies in the field of healthcare has been growing in recent years. As such, routine complete blood count (CBC) and LTCD4 count using flow cytometry (FCM) have been recorded during patient follow-up. CBC and FCM analyses have vital effects on progression monitoring and treatment response assessment in HIV-positive patients [3]. Furthermore, it was found that there were an estimated 39.0 million [33.1–45.7 million] people living with HIV at the end of 2022, two-thirds of whom (25.6 million) are in the WHO African Region [1]. Even though several studies that examined the variation of LTCD4 count, research investigating the prediction of LTCD4 count is relatively limited. Another instrumental factor in the prediction of LTCD4 count is the imbalanced dataset. Addressing the latter presents a challenging procedure that scholars are endeavoring to refine by leveraging various technologies. In pursuit of this objective, Chawla *et al.* (2002) [14] introduced the synthetic minority oversampling technique (SMOTE), recognized as one of the most potent resampling algorithms to solve the imbalance issue by producing synthetic instances from the minor class. Extensive research has proven that SMOTE has a better efficiency than undersampling and oversampling techniques [15, 16]. To our current understanding, limited to no research has been conducted to comprehensively develop the proposed models for predicting LTCD4 counts while incorporating the SMOTE imbalance-aware learning strategy and using CBC and FCM input outcomes across diverse thresholds of total lymphocyte count (TLC). As such, the objectives of this paper are twofold: (1) to ascertain the reliability of TLC as a substitute for LTCD4 count because TLC is easily obtained from CBC by constricting five predictive models of type classification, considering two classes (<200 and ≥200) and (2) to adopt an imbalance-learning strategy based on the SMOTE technique to develop optimized ML models for LTCD4 count.

II. Materials and Methods

A cross-sectional and analytic investigation was undertaken involving 511 patients with HIV infection recruited from the immunology laboratory at Mohammed VI University Hospital (Northern African country) over a 3-year period (2017–2019). The patients were selected from the laboratory database based on their receipt of both CBC and LTCD4 count assessments. Additionally, they were clinically categorized according to the staging criteria of the Centers for Disease Control and Prevention (CDC)[17]. Every patient was diagnosed as having HIV in accordance with WHO recommendations, using combination tests that identify both HIV antibodies and antigens for early and precise diagnosis. The study exclusively considered patients meeting the following criteria: (i) those whose immunological status had been evaluated at the immunology laboratory and (ii) those with comprehensive medical records. Among the 800 patients listed in the hospital registry, 511 were included in our study, whereas 289 were excluded because they did not meet the specified eligibility criteria.

A. Laboratory Testing

A 2-mL sample of venous blood was collected from each patient with HIV on an empty stomach at each time point. Blood cells were analyzed by FCM (FACSCan II, BD Biosciences, San Jose, CA) using a combined CD3/CD4/CD8/CD45 multitest reagent (BD Biosciences, San Jose, CA), allowing the absolute number of lymphocyte subsets to be measured and analyzed. All tests were completed <4 h after venous blood collection. Meanwhile, CBC was measured by routine blood testing using Sysmex XN (Kobe, Japan).

B. Data Preparation Procedure

Patient data essential for the study were reviewed thoroughly, and individuals meeting the eligibility criteria were subsequently included. The dataset was transferred to Microsoft Excel 2010, where it underwent scrutiny and filtration before being exported to statistical analysis software (R software). The analysis primarily concentrated on parameters such as LTCD4 count, CD8 T cell count, CBC, and CDC stage. The normality of continuous variables was evaluated through distribution tests.

C. Methodology

The methodology used in this study conducts a comprehensive evaluation of various lymphocyte count thresholds using ML and deep learning (DP) algorithms. The primary aim is to construct a model for predicting LTCD4 counts by incorporating the most pertinent factors and deploying the most effective "ML/DP" algorithms. The predictive model is a binary classification model designed to ascertain whether an HIV-positive patient will have an LTCD4 count <200. To the best of our knowledge, the exploration of LTCD4 count predictions through the application of diverse algorithms considering different TLC thresholds has not been previously investigated. Consequently, a three-step process is being considered [18]: pretreatment, variable selection, and construction of predictive models (Fig. 1). Preprocessing of the dataset is performed to address the problem of unbalanced class distribution. Variable selection is used to find a set of relevant features to have a robust forecast. Thus, five different models are built: logistic regression (LR), random forest (RF), SVM, ANNs, and naive Bayes (NB). Furthermore, ensuring the absence of data leakage is imperative for maintaining the integrity of predictive modeling in experimental research.

Data leakage occurs when information from the test set influences the training process, leading to overly optimistic model performance [19]. To fortify against such pitfalls, meticulous considerations were taken [20] (Fig. 2).

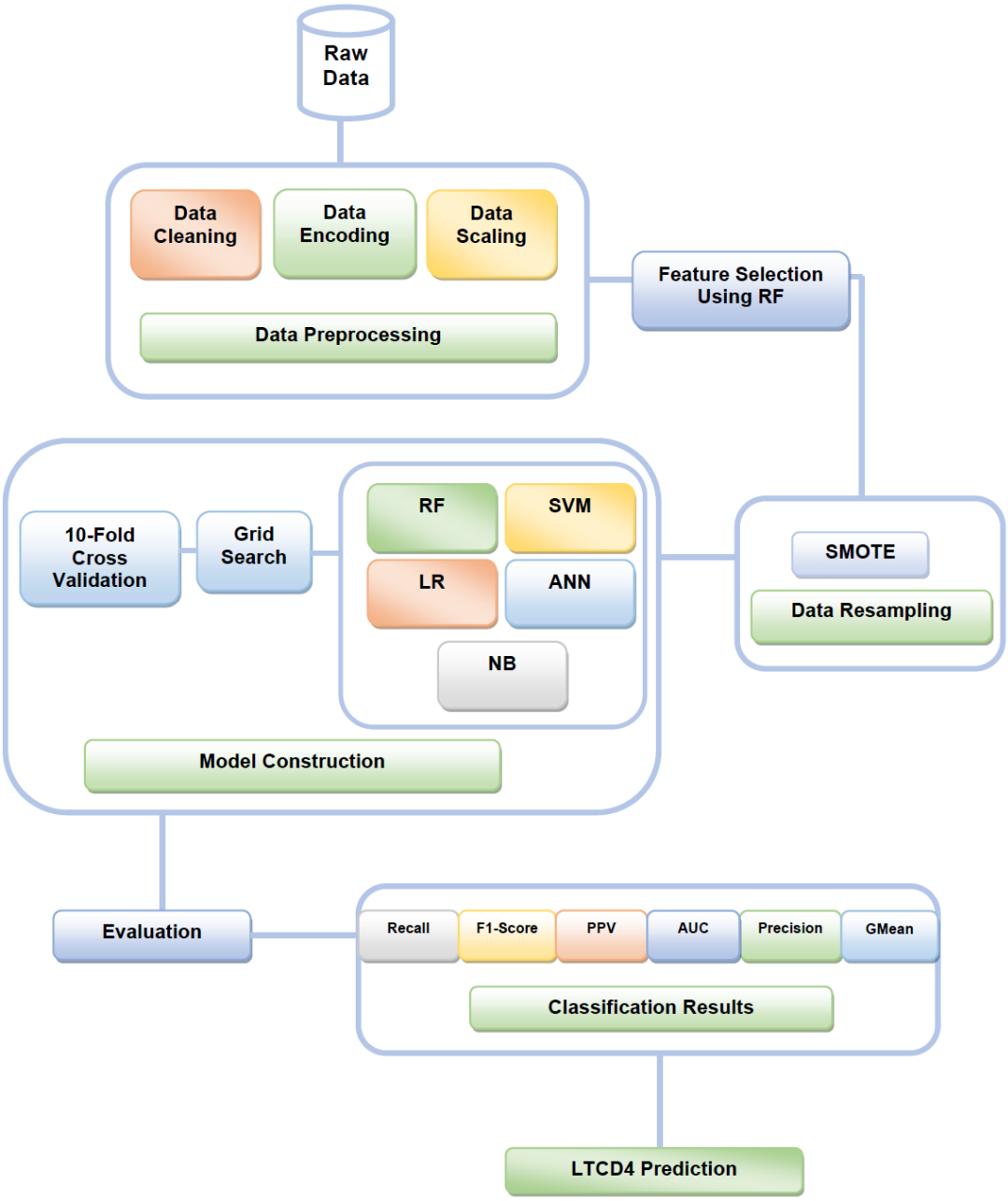

Fig. 1. Overview of the proposed decision framework.

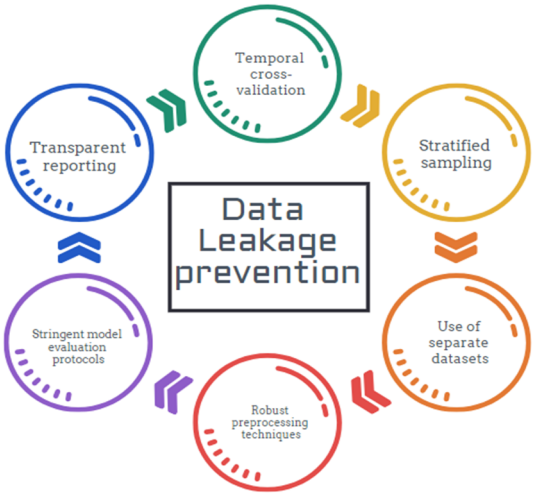

Fig. 2. Data leakage mitigation strategy.

1) Preprocessing

The ratio between the two classes <200 or ≥200 of LTCD4 has generated unbalanced data; such imbalances result in a bias toward the majority class. The classification models prioritize the class with a higher number of observations resulting in an overestimation of this class [21]. Preprocessing aims to solve this problem by balancing the distribution of classes across the dataset. In this paper, SMOTE was adopted [14]. SMOTE creates minority instances based on random intervals between cases.

2) Variable selection

After the preprocessing phase, reducing the number of input variables through variable selection is a crucial step for classification models because some of these characteristics may have no significant effect on the dependent variable.

Thus, there will be an increase in error estimation [22]. Therefore, it is important to inspect our data to determine which variables appear to be strong predictors. The RF model has been widely adopted for variable selection [23]. Thus, nine variables (Fig. 1) were reduced to one (TLC) by the selection process (Fig. 3). The natural history of untreated HIV infection has opposing effects on circulating LTCD4 and CD8 T lymphocytes [24]. Before HIV depletes CD4 cells, circulating CD8 cells will typically increase in response to the infection [22]. In the setting of antiretroviral therapy, some patients will restore CD4 counts and experience a decline in CD8 counts. For other individuals, however, despite suppression of the virus and improvement of CD4 levels, the high levels of circulating CD8 cells are maintained [22]. The imbalance in the game of T cells leads us to suspect that CD8 T lymphocytes may false our results. In our clinical setting, CD8 T lymphocytes are not used for monitoring HIV-positive patients. Consequently, we examined the distribution of both CD8 T lymphocytes and LTCD4 in the remaining 289 patients who did not meet the eligibility criteria but had CD8 T lymphocyte counts documented in their records. Our analysis led to the conclusion that CD8 T lymphocytes do not exert any influence on LTCD4 (Fig. 4).

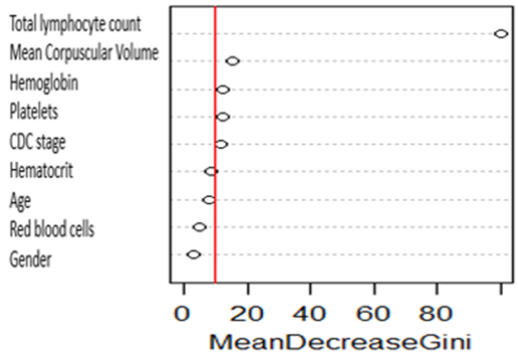

Fig. 3. Variable importance ranking using Gini impurity index.

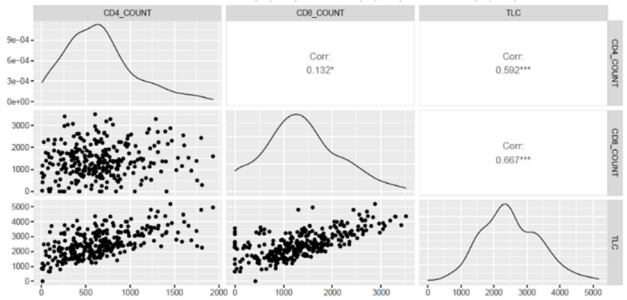

Fig. 4. Distribution and correlation between CD8 T lymphocytes, CD4 T lymphocytes and TLC.

3) Predictive models

a) Logistic regression

In this study, we use LR to forecast LTCD4 by leveraging TLC, formulating the task as a binary outcome prediction endeavor. Acknowledging the pivotal significance of accurate LTCD4 count predictions in clinical decision-making within the field of immunology, our investigation explores the intricate association between TLC and the binary categorization of LTCD4 counts. Using the logistic function, we convert continuous predictors, specifically TLC, into probabilities, refining the model through iterative coefficient adjustments facilitated by advanced optimization algorithms [25]. Coefficients and odds ratios provide measurable indicators of the influence of TLC on the probability of being categorized into specific LTCD4 count groups. Concurrently, performance metrics such as accuracy, sensitivity, and specificity assess the model's effectiveness in predicting binary outcomes [26, 27]. LR can be described as follows:

$$p = \frac{1}{1 + e^{-(\beta 0 + \beta 1 x 1 + \beta 2 x 2 + \cdots + \beta_n x_{n)}}}$$

b) Random forest

The "random forest" algorithm (or RF sometimes also translated as decision tree forest) was proposed by Leo Breiman and Adèle Cutler in 2001 as a statistical prediction or ML algorithm. In the realm of predictive modeling for binary outcome prediction, with a focus on forecasting LTCD4 counts based on TLC, the RF algorithm stands as a methodological cornerstone. This ensemble learning approach, characterized by the aggregation of numerous decision trees, intricately navigates the complexities inherent in immunological data, offering both robustness and interpretability [28]. The methodological rigor of our inquiry initiates with the meticulous curation of a diverse patient cohort, encompassing comprehensive demographic and medical history data [29]. The modeling framework hinges on the dependent variable, signifying binary outcomes corresponding to distinct LTCD4 count categories, and the independent variable, TLC. The RF method operates by generating multiple decision trees, with each tree being trained on a subset of the data and using a random subset of features at each node [30]. This diversity mitigates overfitting and enhances the algorithm's generalization capabilities. The final prediction is a consensus derived from the aggregation of individual tree predictions, resulting in a robust, ensemble-based model [30]. The ensemble nature of RF imparts a layer of stability and reliability to predictions, a critical aspect of clinical decision-making. RF can be described as follows:

$$\hat{y} = \text{mode}(T_1(x), T_2(x), \ldots, T_n(x))$$

c) Support vector machine

Developed by V. N. Vapnik (1995) [31], SVM is a nonprobability binary linear classifier that can be used to solve a classification problem by constructing an optimal hyperplane separation to maximize the margin for SVM to have good generalization capability. SVM is very effective and robust for binary-type classification problems and to perform proportionally or better than other statistical and ML methods [32]. The overarching objective is to contribute to the refined understanding of the intricate relationship between these immunological variables, leveraging SVM's discriminative capabilities. We meticulously define our variables, with the dependent variable representing binary outcomes indicative of distinct LTCD4 count categories, and the independent variable being the TLC. The SVM algorithm, a powerful tool in the domain of ML and statistical modeling, operates by delineating hyperplanes within a high-dimensional feature space [33]. This endeavor aims to maximize the margin between different LTCD4 count categories, effectively capturing the decision boundaries that best segregate these categories. Kernel functions, thoughtfully selected, facilitate nonlinear transformations,

enabling the algorithm to capture intricate relationships that may exist between TLC and LTCD4 count categories. The training process involves identifying support vectors, which are the critical data points influencing the position and orientation of the optimal hyperplane [33]. Subsequently, mathematical optimization techniques are used to iteratively modify hyperplane parameters, aiming for maximum segregation between the binary LTCD4 count categories. In the ensuing scholarly discourse, the results are situated within the wider context of immunology, emphasizing the algorithm's adeptness in discerning complex relationships in data and its sensitivity to optimal feature selection. Recognizing the mathematical elegance of the algorithm and the challenges inherent in its kernelized operations, potential challenges, such as vulnerability to outliers and nuances in parameter tuning, are considered and addressed [34]. SVM can be described as follows:

$$K(x_i, y_i) = \exp(-\gamma ||x_i - y_i||^2)$$
$$f(x) = \text{sgn}(\sum_1^n y_i \alpha_i K(x_i, x_j) + b)$$

d) Artificial neural networks

ANNs are conceptual methodologies inspired by biological neural networks and proven effective and applicable in predicting the relationship between dependent and independent parameters. The forecasting capability of ANNs is significantly influenced by its structure, comprising an input layer, hidden layers, and an output layer. Each layer consists of an interconnected arrangement of units referred to as neurons or nodes, which excel in processing extensive parallel calculations and representing knowledge [35]. This algorithm, inspired by the structure and functioning of the human brain, meticulously navigates the intricacies of immunological data. ANNs comprise interconnected layers of nodes, or neurons, where each connection is assigned a weight [36]. These weights are iteratively adjusted during the training process to optimize the model's performance. The architecture includes an input layer representing features, hidden layers facilitating complex transformations, and an output layer providing predictions [37]. The nonlinear activation functions inherent in each neuron empower the network to capture intricate relationships within the data. Findings derived from the ANNs model reveal insights into the complex and nonlinear dynamics existing between TLC and binary LTCD4 count categories. The model's adaptability allows it to identify patterns that may pose challenges for conventional algorithms. The interpretability concern associated with neural networks is mitigated through the scrutiny of influential neurons and rankings based on feature importance. ANNs can be described as follows:

$$y = f(\sum_j w_{ij} x_j + b)$$

e) Naive Bayes

Operating within the Bayesian paradigm, this algorithm navigates the complexities of immunological data with the precision required for clinical decision-making [38]. NB, rooted in probabilistic reasoning, uses Bayes' theorem to estimate the probability of a patient belonging to a specific LTCD4 count category given their TLC. The "naive" assumption underlying this algorithm is the independence of predictor variables, an oversimplification that nevertheless enables computational efficiency and expeditious model training [39]. In the context of our investigation, this implies assuming independence between TLC and LTCD4 count categories. The algorithm iteratively calculates the probability of a patient falling into each LTCD4 count category, given their TLC. The decision criterion involves selecting the category with the highest probability [39]. This Bayesian approach accommodates incremental updates to probabilities as new data are introduced, fostering adaptability to evolving clinical scenarios. The model's simplicity, interpretability, and computational efficiency make it an attractive option for certain binary classification tasks, albeit with the acknowledgment of its "naive" independence assumption. Although the algorithm's oversimplification may be perceived as a limitation, its robustness and ease of interpretation contribute to its relevance in specific clinical contexts. NB can be described as follows:

$$P(H|E) = \frac{P(E|H) * P(H)}{P(E)}$$

D. Assessment of the Quality of the Model

Various performance measures are frequently used to assess the quality of classification models. They are taken from the contingency table [40]. The contingency table allows the evaluation of the performance by calculating true positive, false positive, false negative, and true negative. In this work, positive predictive value, sensitivity, accuracy, G-mean, F1 score, and area under the curve (AUC) measurements were used. The metrics for model quality assessment are presented and described as follows:

- Recall also called the true-positive rate (TPR) or sensitivity, is defined as the proportion of correctly classified positives. Recall is a particularly substantial metric of classifier performance in this case.

$$Recall = \frac{TP}{TP + FN}$$

- Precision is a measure of accuracy outlining the relevance ratio of the predicted elements, i.e., the percentage of truly predicted events from all predicted events.

$$Precision = \frac{TP}{TP + FP}$$

- G-mean is considered a metric of stability between the correct classification of positive and negative classes viewed independently. It is usually adopted to resist the imbalances in the dataset [41].

$$Gmean = \sqrt{\left(\frac{TP}{TP + FN}\right) \times \left(\frac{TN}{FP + TN}\right)}$$

- F1 score is a highly informative measure as it considers both precision and recall measures, thus taking the class-balance issue into account.

$$F1 \text{ score} = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}}$$

- The positive predictive value (*PPV*) reflects the proportion of positive results. It answers the question, "If

I have a positive test, what is the probability that I have the disease?"

$$PPV = \frac{TP}{PC}$$

- AUC measures the overall performance of the binary classification model. Both TPR and false-positive rates range between 0 and 1. The AUC measures the probability that the model will assign a randomly selected positive instance a higher predicted probability compared with a randomly selected negative instance [42].

The data were partitioned into training and validation sets for model formation and validation 10-fold cross-validation was adopted. This method is recognized for its susceptibility to yield minimal bias and variance in contrast with the other validation methods, including the leave-one-out method [43]. Moreover, k-fold cross-validation has been known to prevent the overfitting issue in the estimation of performance [44]. Data were trained using nine subsets of the input space, and the remaining subset was used to evaluate the performance of predictive learners. The training was repeated 10 times, leaving out one subset that had already been used as a training dataset in the previous training.

III. Results and Discussion

To demonstrate the validity of the classification models, the optimization of the parameters for each classifier was performed. To cope with unbalanced data, SMOTE has been used to rebalance classes; when it comes to unbalanced data, accuracy may suffer because of bias toward the majority class [21]. Therefore, it is essential to select the appropriate measures to evaluate the classifier. Several metrics measure the model's performance, such as AUC, positive predictive value, sensitivity, G-mean, and F1 score. The average prediction of these performances is based on cross-validation, and the results for different classification models are presented in Tables 1–5 and Fig. 5. Based on the results of different models, it appears that the threshold of lymphocyte counts of 1100 had the best performance to predict LTCD4 count. LR (Table 1) demonstrates consistent performance across various lymphocyte counts. The model yields high AUC values (92.01%), signifying commendable discriminative power. However, a notable trade-off between precision and recall is observed, requiring careful consideration of this balance in the clinical context. RF (Table 2) emerges as a robust performer, exhibiting noteworthy performance across diverse lymphocyte counts. Its resilience in handling imbalanced datasets, reflected in high G-mean values (88.18%), positions it as a strong candidate. The model's competitive AUC (93.2%) values underscore its efficacy in discriminating between positive and negative instances. ANNs (Table 3) displays also competitive performance, characterized by high precision (90%) and recall (85%) in most cases. Its adaptability in capturing complex patterns within the data, coupled with high AUC values [93%], highlights its effectiveness in predicting LTCD4 counts. SVM (Table 4) demonstrates strong and consistent performance, particularly excelling in precision (91%) and recall (95%). The high AUC values (97%) affirm its ability to effectively discriminate between positive and negative instances. SVM's efficiency in handling intricate data relationships positions it as a promising choice for LTCD4 count prediction. Although NB (Table 5) exhibits generally good performance, its metrics tend to be slightly lower compared with those of other models.

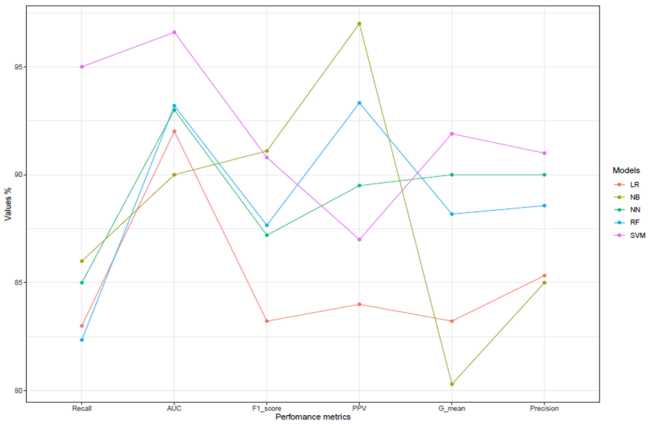

Fig. 5. Performance of all models to predict CD4 T lymphocytes <200/mm3 using the threshold 1100/mm3 of lymphocyte count.

Table 1. Positive predictive value, recall, F1-score, G-mean, AUC and precision using Logistic Regression to predict CD4 T lymphocytes <200/mm3 using different threshold of total lymphocyte count

Models	lymphocyte count	Positive predictive value	Recall	F1-score	G-mean	AUC	Precision
Logistic regression	1000	78.26	72	75	75.06	87	78.5
	1100	84	83	83.22	83.22	92.01	85.33
	1200	85	60	70.03	71.13	86	76
	1300	82.14	61.33	70.22	71.0	88.05	75.6
	1400	62	76	68.26	72.7	84.25	70.5
	1500	80.5	72	76.05	76.11	87.18	79.5
	1600	75.14	70.6	73.10	73	88.52	77
	1700	64.13	78.6	70.65	71.02	81	73

Table 2. Positive predictive value, recall, F1-score, G-mean, AUC and precision using Random Forest to predict CD4 T lymphocytes <200/mm3 using different threshold of total lymphocyte count

Models	lymphocyte count	Positive predictive value	Recall	F1-score	G-mean	AUC	Precision
Random forest	1000	85.71	75	80.17	79.05	91	86.05
	1100	93.33	82.35	87.66	88.18	93.2	88.57
	1200	84.62	73.33	79	81.45	93	88.37
	1300	90.4	82.61	86.45	87.02	92	87.23
	1400	73.6	78	75.7	79.6	92	80
	1500	89.74	85	87.20	88.57	90	89.13
	1600	77.6	75	76.37	75	89	75
	1700	81	70	75.26	77	80	79.31

Table 3. Positive predictive value, recall, F1-score, G-mean, AUC and precision using Neural Network to predict CD4 T lymphocytes <200/mm3 using different threshold of total lymphocyte count

Models	lymphocyte count	Positive predictive value	Recall	F1-score	G-mean	AUC	Precision
Neural network	1000	80.3	81	80.6	83.5	82.4	76
	1100	89.50	85	87.2	90	93	90
	1200	80.7	85.1	80.31	83	82	76.3
	1300	80	84.21	82.05	85.64	85.6	78
	1400	87.50	77.78	82.53	82.73	85.2	82.7
	1500	83.33	77.78	75.67	78.88	86.3	79.07
	1600	79.3	71.43	77	79.5	75.54	81
	1700	73.5	75	74.24	74.2	70	76.74

Table 4. Positive predictive value, recall, F1-score, G-mean, AUC and precision using Support Vector Machine to predict CD4 T lymphocytes <200/mm3 using different threshold of total lymphocyte count

Models	lymphocyte count	Positive predictive value	Recall	F1-score	G-mean	AUC	Precision
SVM	1000	82	85	83.4	81.5	95	82
	1100	87	95	91	92	97	91
	1200	80	85	84.2	84	98	84
	1300	77	81	79	81.4	95.9	82
	1400	81	83	82	84	95	84
	1500	85	68	75.5	82.04	96.3	95
	1600	89	84	86.4	88.3	95.07	90
	1700	81	93	86	86.2	96	86

Table 5. Positive predictive value, recall, F1-score, G-mean, AUC and precision using Naïve Baiyes to predict CD4 T lymphocytes <200/mm3 using different threshold of total lymphocyte count

Models	lymphocyte count	Positive predictive value	Recall	F1-score	G-mean	AUC	Precision
Naive baiyes	1000	80	79	79.5	82	89	82
	1100	97	86	89	80.3	90	85
	1200	82	86	84	86	91	86
	1300	85	86	85.4	87	93	87
	1400	78	73	75.17	79	85	80
	1500	77	81	79	81.5	89	82
	1600	84	75	79.24	81.7	87.8	83
	1700	73	75	73.9	76.9	82	77

Average life expectancy has increased significantly over the past century as a result of advances in technology [32]. Healthcare is one of the largest industries in the world that can capitalize on the advances in technology [45]. The term "machine learning" is used to describe various statistical techniques that allow computers to learn from experiences without being explicitly programmed. ML has many applications in healthcare systems, such as big data tools and mandatory procedures, like electronic health records [13]. ML enhances the quality of automation and smart decision-making in primary and tertiary care and public health systems, which could be the biggest effect of ML tools and can improve the quality of life of billions of people worldwide [46]. As cutting-edge technology grows in popularity in the healthcare industry, it produces high-performance computing, fast, reliable, and able to process large and complex data. Automated ML helps health professionals deliver high-quality patient care, with more efficiency [47], and to make better-informed decisions. The use of ML in healthcare settings allows physicians to recommend treatments or even reduce drastically its severity. It is especially important in patients with HIV infection, where lymphocyte depletion, mainly of the LTCD4 cell subset due to immunodeficiency, has been recognized as a hallmark for monitoring patients with HIV infection [48]. Enhanced accessibility to monitoring tools, particularly LTCD4 count and viral load testing for individuals undergoing antiretroviral therapy, holds significant importance [49, 50]. Nonetheless, in low-income countries, restricted availability of advanced immunoassays necessitates reliance on clinical staging for patient follow-up. Ideally, the WHO advocates for the routine adoption of combined immunological and virologic monitoring in all patients with HIV infection [51]. However, the analysis of viral load and LTCD4 count requires not only sophisticated equipment but also highly skilled personnel (WHO, 2006) [52]. This gives rise to substantial financial constraints, especially in settings with limited resources [53]. Consequently, there is a pressing need for more cost-effective alternative approaches. Thus, a threshold analysis was performed in this study using ML and DL algorithms to determine the ability of TLC to predict LTCD4 count. The results of this study demonstrate that near-accurate predicted levels were achieved using SVM, ANNs, LR, RF, and NB with an "1100" lymphocyte count as a threshold. SVM outperforms all algorithms with a classification precision of 91%. However, high-precision algorithms are not always synonymous with better performance. Therefore, the choice of algorithm should consider the specific context and requirements of the application. High-precision algorithms may not always be the best choice if they come with significant computational costs or lack interpretability [54, 55]. A balanced approach that takes into account multiple performance metrics, computational efficiency, and model interpretability is essential for selecting the most appropriate algorithm for a given task. This comprehensive evaluation framework ensures that the chosen algorithm delivers optimal performance across all relevant dimensions, rather than excelling in just one. This balanced view is supported by several corroborative studies, which affirm the reliability of these principles and the importance of considering a broad range of performance metrics when evaluating algorithmic performance [56, 57]. Other metrics such as AUC are considered robust measures for imbalanced data [58]. AUC is a pivotal metric, providing a comprehensive assessment of

discrimination ability. The AUC values for five different algorithms with an "1100" lymphocyte count as a threshold are as follows: LR (92.01%), RF (93.2%), ANNs (93%), SVM (97%), and NB (90%). These percentages represent the models' ability to distinguish between positive and negative instances, with higher values indicating superior discriminatory power. The practical implications of a 5% difference in AUC between SVM and LR models are noteworthy. In predictive modeling, a higher AUC indicates improved discrimination, indicating that the SVM model, with its 97% AUC, exhibits superior performance compared with the LR model, which has an AUC of 92.01%. However, it is crucial to interpret these differences contextually. Although the SVM model excels in discrimination, achieving a 5% higher AUC, the LR model's unique strength lies in its interpretability. LR allows for the calculation of marginal effects for each variable, providing insights into how changes in predictors affect the probability of a positive outcome. This interpretability is particularly valuable in real-world applications and decision-making scenarios, offering a transparent understanding of the contribution of individual features to the predicted outcome. In the context of HIV prediction tasks, this 5% difference in AUC could translate to a more nuanced understanding of the discriminative capabilities of the SVM model over LR. Although SVM excels in capturing complex, nonlinear relationships, LR's interpretability allows for the identification of specific factors contributing to the likelihood of HIV prediction. This insight is crucial in a medical context, where understanding the influence of various variables on the prediction can inform targeted interventions and decision-making processes. In other words, the 5% difference in AUC between SVM and LR models signifies a trade-off between discriminative power and interpretability. The SVM model demonstrates superior discrimination, whereas the LR model offers a transparent understanding of variable contributions. Depending on the specific goals of the HIV prediction task, stakeholders may prioritize either model based on their distinct strengths – the nuanced interpretability of LR or the discriminative capability of SVM. Other evaluation metrics such as recall, positive predictive value, G-mean, and F1 score showed superior performance for SVM over the rest of the algorithms used in this paper. Hence, SVM can be used to develop cost-effective web-based prediction models to help forecast patients' future LTCD4 count changes without performing costly LTCD4 immunoassay. Actually, in many developing countries, the main challenges are the scarcity of expensive machines such as FCM, frequent machine breakdowns, lack of timely and proper maintenance, and lack of reagents [59–61]. With large datasets available on HIV/AIDS, the use of data-based predictive models is increasingly being explored. As a result, the use of these models could be an alternative to overcoming LTCD4 account challenges in resource-constrained contexts. Furthermore, a study that used decision trees and RF algorithms revealed that the TLC, hemoglobin, and total platelet counts have significant prognostic value for monitoring HIV/AIDS clinical progression [62]. Comparable predictive performance accuracies of data mining algorithms were reported by previous studies [63, 64]. An RF algorithm was applied to HIV FCM data to predict LTCD4 immune reconstitution outcomes [64]. Wang *et al.* compared RF and SVM algorithms to accurately predict the virologic response of patients with antiretroviral treatment [65]. RF has identified that the LTCD4 count cutoff value of 400 cells/μL was a power classifier [65]. A study from the UK reported that RF and SVM models can produce predictions of virological response to HIV treatment [65]. In light of our findings and the existing body of research, ML algorithms emerge as a compelling solution for predicting LTCD4 counts in resource-constrained settings. It is recommended that governments leverage big data methodologies to implement predictive models for LTCD4 counts. The insights gained from this study not only advance the field of HIV prediction but also offer valuable methodologies that can be translated to other domains, including sustainable energy management. For instance, in areas prone to natural disasters or those undergoing rapid environmental changes, accurately predicting resource needs and understanding the effect of these changes is critical. Our approach can be tailored to address specific local conditions, providing targeted solutions that improve resilience and sustainability [66]. By integrating ML techniques into the design of energy optimization strategies, we can enhance the adaptability and responsiveness of these strategies to local ecological disturbances [67, 68]. Nevertheless, our study has several limitations. The dataset has a low number of observations. ML/DP algorithms usually work best when used with a large number of data points and variables [69]. Therefore, one of the limitations of this study is that it did not include important variables that are known to be potentially associated with the LTCD4 count such as the viral load [63], types of opportunistic infections, and nutritional status of the patient. Future studies need to consider including these variables if a more robust classification result is to be achieved. However, because of resource constraints, viral load is unavailable in many developing countries [70]. Hence, HIV care in these countries is dependent on periodic LTCD4 count and WHO clinical staging. Future studies also need to test more algorithms, preferably ones that can predict the absolute count.

IV. Conclusion

The data of this study allowed defining a threshold of lymphocyte counts of 1100 to predict the LTCD4 count using machine learning models: Random Forest, Naive Bayes, Support Vector Machines, Logistic Regression, and Artificial Neural Networks. In accordance with similar studies, the results of this study indicate that machine learning algorithms have the potential to replace costly immunoassays in resource-limited regions by building web-based system prediction that can help providers monitor patients' health status and make recommendations for better management, prognosis, and resource allocation.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Saad Lamjadli, Oumayma Ouedrhiri: Conceptualization, Methodology, Software; Saad Lamjadli, Ikram Souli, Safa Machraoui: Data curation, Writing-Original draft,

Preparation; Zouhair Elamrani Abou Elassad, Oumayma Banouar, Said Raghay, Brahim Admou: Visualization, Investigation; Moulay Yassine Belghali, Raja Hazime, Noura Tassi, Said Raghay, Brahim Admou: Supervision; Oumayma Ouedrhiri, Oumayma Banouar, Said Raghay, Brahim Admou: Validation.

REFERENCES

[1] WHO. HIV and AIDS. (2023). [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/hiv-aids

[2] WHO, "Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection Switzerland", WHO, June 2013. 2013.

[3] F. N. Meintjes, A. Pozniak *et al.,* "The future role of CD4 cell count for monitoring antiretroviral therapy," *Lancet Infect Dis 2015*, vol. 15, no. 2, pp. 241–247, 2015.

[4] H. Birhan, A. Seyoum, K. Derebe, S. Muche, M. Wale, and S. Sisay, "Joint clinical and socio-demographic determinants of CD4 cell count and body weight in HIV/TB co-infected adult patients on HAART," *Scientific African*, 2022.

[5] H. Mousannif, A. Karkouch *et al.*, "Engineering applications of artificial intelligence the application of machine learning techniques for driving behavior analysis: A conceptual framework and a systematic literature review," *Eng Appl Artificial Intell*., vol. 87, 2019.

[6] B. Li, M. Li, Y. Song *et al*, "Construction of machine learning models to predict changes in immune function using clinical monitoring indices in HIV/AIDS patients after 9.9-Years of antiretroviral therapy in Yunnan, China," *Front Cell Infect Microbiol.*, May 12, 2022.

[7] R. Yu. M. Abdel-Aty, "Utilizing support vector machine in realtime crash risk evaluation," *Acc Anal Prevent.*, vol. 51, pp. 252–259, 2013. doi:10. 1016/j.aap.2012.11.027

[8] S. Basu, S. Mukhopadhyay, M. Karki, R. DiBiano *et al.*, "Deep neural networks for texture classification—A theoretical analysis," *Neural Networks*, vol. 97, no 173–82, 2018. doi: 10.1016/j.neunet.2017.10.001

[9] C. W. Hsu, C. C. Chang *et al*., A Practical Guide to Support Vector Classification, 2003.

[10] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *J Mach Learn Res*., vol. 15, no. 1929–1958, 2014.

[11] R. Kohavi. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.48.529

[12] R. A. Hughes, M. T. May, K. Tilling, N. Taylor *et al.*, "Long terms trends in CD4+ cell counts, CD8+ cell counts, and the CD4+: CD8+ ratio," *AIDS*, vol. 32, 10, pp. 1361-1367, 2018. doi: 10.1097/QAD.0000000000001848.

[13] A. Ahlström, L. H. Omland et al,. "Algorithmic prediction of HIV status using nation-wide electronic registry data," *EClinicalMedicine*, vol. 17, 2019. doi: 10.1016/j.eclinm.2019.10.016. PMID: 31891137; PMCID: PMC6933258. 2019.

[14] N. V. Chawla, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321–357, 2002. doi:10.1613/jair.953. 2002.

[15] R. Batuwita and V. Palade, "Class imbalance learning methods for support vector machines," *Imbalanced Learning*, Hoboken, NJ: John Wiley & Sons, Inc. pp. 83–99. 2013.

[16] H. M. Nguyen. E. W. Cooper, and K. Kamei, "Borderline over-sampling for imbalanced data classification," *International Journal of Knowledge Engineering and Soft Data Paradigms*, vol. 3, no. 1, p. 4, 2011. doi:10. 1504/IJKESDP.2011.039875. 2011.

[17] CDC. (2024). [Online]. Availale: www.cdc.gov/hiv

[18] Z. E.A. Elassad, H. Mousannif, and H. A. Moatassime, "Class-imbalanced crash prediction based on real-time traffic and weather data: A driving simulator study," *Traffic Injury Prevention*, 2020. DOI: 10.1080/15389588.2020.1723794

[19] S. Kapoor and A. Narayanan, "Leakage and the reproducibility crisis in machine-learning-based science," *Patterns*, vol. 4, no. 9, p. 100804, 2023. doi: 10.1016/j.patter.2023.100804

[20] Q. Dong, "Leakage Prediction in machine learning models when using data from sports wearable sensors," *Comput Intell Neurosci*., 2022. doi: 10.1155/2022/5314671

[21] A. Fernandez, "Hierarchical fuzzy rule-based classification systems with genetic rule selection for imbalanced data-sets," *Int. J. Approx. Reason.*, 2009. doi.org/10.1016/j.ijar.2008.11.004

[22] M. Helleberg. G. Kronborg, H. Ullum *et al.*, "Course and clinical significance of CD8+ T-Cell counts in a large cohort of HIV-infected individuals," *J Infect Dis.*, 2015, vol. 211, no. 11, pp. 1726-34. doi: 10.1093/infdis/jiu669. Epub 2014 Dec 8. PMID: 25489001. 2014.

[23] A. Cutler, R. Cutler, and J. R. Stevens, "Random Forests," *Ensemble Machine Learning*, Springer, New York, NY. https://doi.org/10.1007/978-1-4419-9326-7_5. 2012.

[24] J. B. Margolick. A. Muñoz, A. D. Donnenberg, L. P. Park, N. Galai, J. V. Giorgi, M. R. O'Gorman, and J. Ferbas, "Failure of T-cell homeostasis preceding AIDS in HIV-1 infection," *The Multicenter AIDS Cohort Study. Nat Med*., vol. 1, no. 7, pp. 674-80, 1995. doi: 10.1038/nm0795-674. PMID: 7585150. 1995.

[25] S. Sperandei, "Understanding logistic regression analysis," *Biochem Med (Zagreb)*, vol. 24, no. 1, pp. 12-18, 2014. doi: 10.11613/BM.2014.003. PMID: 24627710; PMCID: PMC3936971. 2014.

[26] W. Zhang, G. Huang, K. Zheng, J. Lin, S. Hu *et al.,* "Application of logistic regression and machine learning methods for idiopathic inflammatory myopathies malignancy prediction," *Clin Exp Rheumatol*. vol. 41, no. 2, pp. 330-339, 2023,. doi: 10.55563/clinexprheumatol/8ievtq. Epub 2023 Mar 1. PMID: 36861746. 2023.

[27] X. Song, X. Liu, F. Liu, and C. Wang, "Comparison of machine learning and logistic regression models in predicting acute kidney injury: A systematic review and meta-analysis," *Int J Med Inform*., 2021, vol. 151, p. 104484, 2021.

[28] S. Kamal. J. Urata, M. Cavassini, H. Liu, R. Kouyos, O. Bugnon, W. Wang, and M. P. Schneider, "Random Forest machine learning algorithm predicts virologic outcomes among HIV infected adults in Lausanne, Switzerland using electronically monitored combined antiretroviral treatment adherence," *AIDS Care.*, vol. 33, no. 4, pp. 530-536, 2021. doi: 10.1080/09540121.2020.1751045. Epub 2020 Apr 8. PMID: 32266825. 2021.

[29] D. N. Mamo, T. M. Yilma, M. Fekadie, Y. Sebastian, T. Bizuayehu, M. S. Melaku, and A. D. Walle. "Machine learning to predict virological failure among HIV patients on antiretroviral therapy in the University of Gondar Comprehensive and Specialized Hospital, in Amhara Region, Ethiopia, 2022," *BMC Med Inform Decis Mak.*, vol. 23, no. 1, p. 75, 2023. doi: 10.1186/s12911-023-02167-7

[30] J. He. J. Li, S. Jiang, W. Cheng, J. Jiang, Y. Xu, J. Yang, X. Zhou, C. Chai, and C. Wu, "Application of machine learning algorithms in predicting HIV infection among men who have sex with men: Model development and validation," *Front Public Health* vol. 10, 2022. doi: 10.3389/fpubh.2022.967681. PMID: 36091522; PMCID: PMC9452878. 2022.

[31] V. N. Vapnik, *The Nature of Statistical Learning Theory*, New York: Springer, 1995.

[32] G. Nalinipriya. G. Nalinipriya, P. Priyadarshini, S. S. Puja, K. R. R. Baymax, "A smart healthcare system provides services to millennials using machine learning technique," in *Proc. 2019 International Conference on Smart Structures and Systems (ICSSS)*, 2019.

[33] S. Huang, N. Cai, P. P. Pacheco, S. Narrandes, Y. Wang, and W. Xu, "Applications of Support Vector Machine (SVM) Learning in cancer genomics," *Cancer Genomics Proteomics*, vol. 15, no. 1, pp. 41-51, 2018. doi: 10.21873/cgp.20063. PMID: 29275361; PMCID: PMC5822181. 2018.

[34] A. Ben-Hur and J. Weston, "A user's guide to support vector machines," *Methods Mol Biol.* vol. 609, pp. 223-39, 2010,. doi: 10.1007/978-1-60327-241-4_13. PMID: 20221922

[35] J. Zou, Y. Han, and S. S. So, "Overview of artificial neural networks," *Methods Mol Biol.*, vol. 458, pp. 15-23, 2008. doi: 10.1007/978-1-60327-101-1_2

[36] Z. Zhang, "A gentle introduction to artificial neural networks," *Ann Transl Med*., vol. 4, no. 19, pp. 370, 2016. doi: 10.21037/atm.2016.06.20

[37] S. H. Han, K. W. Kim, S. Kim, and Y. C. Youn. "Artificial Neural network: Understanding the basic concepts without mathematics," *Dement Neurocogn Disord*., vol. 17, no. 3, pp. 83-89, 2018. doi: 10.12779/dnd.2018.17.3.83

[38] D. Musgrove, J. Wolfson, S. Bandyopadhyay *et al.*, "A Naive Bayes machine learning approach to risk prediction using censored, time-to-event data," *Stat Med*., vol. 34, no. 21, pp. 2941-57. doi: 10.1002/sim.6526, 2015.

[39] M. Langarizadeh and F. Moghbeli, "Applying Naive Bayesian networks to disease prediction: A systematic review," *Acta Inform Med.*, vol. 24, no. 5, pp. 364-369, 2016. doi: 10.5455/aim.2016.24.364-369

[40] F. D. Kakhki, S. A. Freeman, and G. A. Mosher, "Evaluating machine learning performance in predicting injury severity in agribusiness industries," *Safety Science*, vol. 117, pp. 257–62, 2019. doi: 10.1016/j.ssci.2019.04.026

[41] M. Kubat. M. Kubat, and S. Matwin, "Addressing the curse of imbalanced training sets: One-sided selection," in *Proc. Fourteenth Int. Conf. Mach. Learn*., 1997, pp. 179–186.
[42] P. A. Bradley, "The use of the area under the ROC curve in the evaluation of machine learning algorithms," *Pattern Recognition*, vol. 30, no. 7, pp. 1145–1159, 1997.
[43] R. Kohavi and R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," *Proceedings IJCAI-95*, pp. 1137–1143, 1995.
[44] F. Nafiah, A. Sophian, M. R. Khan, and I. M. Z. Abidin, "Quantitative evaluation of crack depths and angles for pulsed eddy current non-destructive testing," *NDT E Int.*, vol. 2, pp. 180–188, 2019.
[45] A. Abdelaziz, M. Elhoseny, A. S. Salama, and A. M. Riad, "A machine learning model for improving healthcare services on cloud computing environment," *Measurement*, vol. 119. pp. 117–128, 2018.
[46] A. Tucker. Z. Wang, Y. Rotalinti, and P. Myles, "Generating high-fidelity synthetic patient data for assessing machine learning healthcare software," *NPJ Digit Med*., vol. 3, no. 1, p. 147, 2020. doi: 10.1038/s41746-020-00353-9. PMID: 33299100; PMCID: PMC7653933.
[47] H. Habehh and S. Gohel, "Machine learning in healthcare," *Curr Genomics.*, vol. 22, no. 4, pp. 291-300, 2021. doi: 10.2174/1389202922666210705124359
[48] D. Daka and E. Loha, "Relationship between total lymphocyte count (TLC) and CD4 count among peoples living with HIV, Southern Ethiopia: A retrospective evaluation," *AIDS Res Ther.*, vol. 5, no. 26, 2008. doi: 10.1186/1742-6405-5-26
[49] P. H. Kilmarx and R. Simbi, "Progress and challenges in scaling up laboratory monitoring of HIV treatment," *PLoS Med*., vol. 13, no. 8, p. e1002089, 2016. doi: 10.1371/journal.pmed.1002089
[50] S. P. Johnstone-Robertson, J. Hargrove, and B. G. Williams, "Antiretroviral therapy initiated soon after HIV diagnosis as standard care: potential to save lives?" *HIV AIDS (Auckl).*, vol. 3, pp. 9-17, 2011. doi: 10.2147/HIV.S7278
[51] UNAIDS, WHO AIDS Epidemic Update, 2007.
[52] World Health Organization, *Antiretroviral therapy for HIV Infection in Adults and Adolescents: Recommendations for a Public Health Approach*, 2006.
[53] S. Crowe, S. Turnbull, R. Oelrichs, and A. Dunne. "Monitoring of human immunodeficiency virus infection in resource-constrained countries," *Clin Infect Dis*., vol. 37, pp. S25-35, 2003. doi: 10.1086/375369
[54] E. Bazmi, B. Behnoush, S. N. Hashemi *et al.*, "Seizure prediction model in acute tramadol poisoning; A derivation and validation study," *Arch Acad Emerg Med*., vol. 8, no. 1, p. e59, 2020. PMID: 32613201; PMCID: PMC7305636.
[55] B. Behnoush, E. Bazmi, S. H. Nazari *et al*., "Machine learning algorithms to predict seizure due to acute tramadol poisoning," *Hum Exp Toxicol*., vol. 40, no. 8, pp. 1225-1233, 2021. doi: 10.1177/0960327121991910
[56] N. Kitcharanant, P. Chotiyarnwong, T. Tanphiriyakun *et al.*, "Development and internal validation of a machine-learning-developed model for predicting 1-year mortality after fragility hip fracture," *BMC Geriatr*., vol. 22, no. 1, p. 451, 2022. doi: 10.1186/s12877-022-03152-x
[57] X. Liang, X. Li, G. Li *et al.*, "A machine learning approach to predicting vascular calcification risk of type 2 diabetes: A retrospective study," *Clin Cardiol.*, vol. 47, no. 4, p. e24264, 2024. doi: 10.1002/clc.24264
[58] X. R. Chao, "An efficiency curve for evaluating imbalanced classifiers considering intrinsic data characteristics: Experimental analysis," *Information Sciences*, vol. 608, pp. 1131–1156, 2022.
[59] V. Habiyambere, N. Ford, D. Low-Beer, J. Nkengasong, A. Sands, M. Pérez González, P. Fernandes, and E. Milgotina, "Availability and use of HIV monitoring and early infant diagnosis technologies in WHO member states in 2011-2013: Analysis of annual surveys at the facility level," *PLoS Med*., vol. 13, no. 8, p. e1002088, 2016. doi: 10.1371/journal.pmed.1002088
[60] T. Peter, C. Zeh, Z. Katz, A. Elbireer *et al*., "Scaling up HIV viral load-lessons from the large-scale implementation of HIV early infant diagnosis and CD4 testing," *J Int AIDS Soc.*, 2017. doi: 10.1002/jia2.25008. PMID: 29130601
[61] F. Bianchi, J. Cohn, E. Sacks, R. Bailey, J. F. Lemaire *et al*., "Evaluation of a routine point-of-care intervention for early infant diagnosis of HIV: An observational study in eight African countries," *Lancet HIV*, no. 6, pp. e373-e381, 2019. doi: 10.1016/S2352-3018(19)30033-5
[62] H. Rafatpanah, L. Essmailian, M. R. Hedayati-Moghaddam, R. Vakili, M. Norouzi, M. R. Sarvghad, A.M. Hosseinpour, H. Sharebiani, and S. A. Rezaee, "Evaluation of non-viral surrogate markers as predictive indicators for monitoring progression of human immunodeficiency virus infection: An eight-year analysis in a regional center," *Jpn J Infect Dis.*, vol. 69, no. 1, pp. 39-44, 2016. doi: 10.7883/yoken.JJID.2014.261
[63] J. Pérez-Santiago, D. Ouchi, V. Urrea, J. Carrillo, C. Cabrera, J. Villà-Freixa, J. Puig, R. Paredes, E. Negredo, B. Clotet, M. Massanella, and J. Blanco, "Antiretroviral therapy suppressed participants with low CD4+ T-cell counts segregate according to opposite immunological phenotypes," *AIDS.*, vol. 30, no. 15, pp. 2275-2287, 2016. doi: 10.1097/QAD.0000000000001205
[64] M. Eliot, C. Firnhaber, W. Stevens, D. K. Glencross, I. Sanne, L. J. Montaner, and A. S. Foulkes, "Tree-based methods for discovery of association between flow cytometry data and clinical endpoints," *Adv Bioinformatics*, p. 235320, 2009. doi: 10.1155/2009/235320.
[65] D. Wang, B. Larder, A. Revell *et al*., "Comparison of three computational modelling methods for the prediction of virological response to combination HIV therapy," *Artif Intell Med.*, vol. 47, no. 1, pp. 63-74, 2009. doi: 10.1016/j.artmed.2009.05.002
[66] S. Koley, "Contemporary practices in groundwater arsenic remediation and wastewater management in West Bengal, India: A systematic review," *International Journal of Advanced Technology and Engineering Exploration*, 2021.
[67] Q. Yu, W. Ji, L. Prihodko *et al*., "Study becomes insight: Ecological learning from machine learning," *Methods Ecol Evol.*, vol. 12, no. 11, pp. 2117-2128, 2021. doi: 10.1111/2041-210X.13686
[68] T. Welchowski, K. O. Maloney, R. Mitchell *et al*., "Techniques to improve ecological interpretability of black-box machine learning models," *J Agric Biol Environ Stat*., vol. 27, pp. 175-197, 2021. doi: 10.1007/s13253-021-00479-7
[69] M. Z. F. Khairuddin, P. L.Hui *et al*., "Occupational injury risk mitigation: Machine learning approach and feature optimization for smart workplace surveillance," *Int J Environ Res Public Health*, 2022, vol. 19, no. 21, p. 13962. doi: 10.3390/ijerph192113962. PMID: 36360843; PMCID: PMC9653932. 2022.
[70] WHO, "Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection: Recommendations for a public health approach," *National Library of Medicine*, Geneva: WHO, 2013.