algorithms

Systematic Review

Utilization of Machine Learning Algorithms for the Strengthening of HIV Testing: A Systematic Review

Musa Jaiteh [1,*], Edith Phalane [1], Yegnanew A. Shiferaw [2], Karen Alida Voet [3] and Refilwe Nancy Phaswana-Mafuya [1]

[1] Faculty of Health Sciences, South African Medical Research Council/University of Johannesburg (SAMRC/UJ) Pan African Centre for Epidemics Research (PACER) Extramural Unit, Johannesburg 2006, South Africa; edithp@uj.ac.za (E.P.); refilwep@uj.ac.za (R.N.P.-M.)
[2] Department of Statistics, Faculty of Science, University of Johannesburg, Johannesburg 2092, South Africa; yegnanews@uj.ac.za
[3] Department of International Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD 21205, USA; kvoet1@jhu.edu
* Correspondence: mjaiteh1993@gmail.com; Tel.: +27-791850627

Abstract: Several machine learning (ML) techniques have demonstrated efficacy in precisely forecasting HIV risk and identifying the most eligible individuals for HIV testing in various countries. Nevertheless, there is a data gap on the utility of ML algorithms in strengthening HIV testing worldwide. This systematic review aimed to evaluate how effectively ML algorithms can enhance the efficiency and accuracy of HIV testing interventions and to identify key outcomes, successes, gaps, opportunities, and limitations in their implementation. This review was guided by the Preferred Reporting Items for Systematic Reviews and Meta-Analysis guidelines. A comprehensive literature search was conducted via PubMed, Google Scholar, Web of Science, Science Direct, Scopus, and Gale OneFile databases. Out of the 845 identified articles, 51 studies were eligible. More than 75% of the articles included in this review were conducted in the Americas and various parts of Sub-Saharan Africa, and a few were from Europe, Asia, and Australia. The most common algorithms applied were logistic regression, deep learning, support vector machine, random forest, extreme gradient booster, decision tree, and the least absolute shrinkage selection operator model. The findings demonstrate that ML techniques exhibit higher accuracy in predicting HIV risk/testing compared to traditional approaches. Machine learning models enhance early prediction of HIV transmission, facilitate viable testing strategies to improve the efficiency of testing services, and optimize resource allocation, ultimately leading to improved HIV testing. This review points to the positive impact of ML in enhancing early prediction of HIV spread, optimizing HIV testing approaches, improving efficiency, and eventually enhancing the accuracy of HIV diagnosis. We strongly recommend the integration of ML into HIV testing programs for efficient and accurate HIV testing.

Keywords: machine learning; algorithms; HIV testing; predictive modeling; support vector machine; XGBoost; LASSO; random forest; deep learning; efficient; accurate

Citation: Jaiteh, M.; Phalane, E.; Shiferaw, Y.A.; Voet, K.A.; Phaswana-Mafuya, R.N. Utilization of Machine Learning Algorithms for the Strengthening of HIV Testing: A Systematic Review. *Algorithms* **2024**, *17*, 362. https://doi.org/10.3390/a17080362

Academic Editor: Frank Werner

Received: 27 May 2024
Revised: 1 August 2024
Accepted: 13 August 2024
Published: 17 August 2024

1. Introduction

For the past four decades, human immunodeficiency virus/acquired immunodeficiency syndrome (HIV/AIDS) has been a devastating burden to populations worldwide. According to the recent World Health Organization (WHO) estimates, HIV/AIDS has infected 85.6 million individuals and killed 40.4 million people since the epidemic began [1]. By the end of 2022, at least 39.0 million people were living with HIV (PLHIV), of which 25.6 million lived in Sub-Saharan Africa (SSA) [1]. Despite the global attempts to address HIV/AIDS, high HIV incidence and reluctance to test for HIV persist, particularly in South and Eastern Africa, which bear the heaviest HIV burden [1,2]. Accelerated action to end the AIDS epidemic by 2030, also known as the Fast-Track (95-95-95), was developed in 2020 by the Joint United Nations Program on HIV/AIDS (UNAIDS) [3]. This HIV/AIDS response established an interim target for 2025, proposing that 95% of PLHIV know their status, 95% of those who know their status are linked to treatment, and 95% of those who are on treatment are virally suppressed [3]. If achieved, this would mark a magnificent milestone in the history of epidemiology and reshape healthcare sectors for the greater good.

As part of the cross-sectoral efforts to meet the UNAIDS 2030 target, the WHO developed a consolidated framework for HIV prevention and care, with a specific emphasis on prioritizing new HIV testing strategies [2]. HIV testing serves as the gateway to realizing the UNAIDS Fast-Track goal because being tested is the only way to be catered for in a specific HIV intervention. However, the recent UNAIDS report indicated that only 86% of PLHIV are aware of their status, while 90% of PLHIV are currently on treatment [4].

Although a significant number of PLHIV have been accurately diagnosed and linked to treatment through conventional counseling and testing approaches over the years, these methods still fall short of adequately serving certain groups [5,6]. Key populations (KPs), who are often the main drivers of HIV transmission, frequently face stigma and discomfort when accessing public healthcare services, which has led the WHO to recommend HIV self-testing (HIVST) [7]. Furthermore, traditional statistical models, while useful, often struggle to predict HIV risk when applied to complex datasets [6].

As we approach the 2030 agenda, conventional HIV testing methods alone seem inadequate for the ambitious goal of ending HIV as an epidemic [6]. Limitations of traditional HIV testing, including issues related to accessibility, acceptability, and privacy, have become more apparent. As a result, their effectiveness in diagnosing high-risk populations remains questionable [5–7]. Moreover, the impacts of COVID-19 on the global economy are still fresh, poising resource constraints on the implementation of universal screening for HIV, especially in developing countries [8,9]. Therefore, these issues, coupled with emerging epidemics, demand the integration of more innovative HIV testing approaches such as machine learning (ML) [10].

Machine learning involves the use of computational and statistical algorithms that learn from complex data to identify hidden patterns and associations, improving the efficacy of predictions and the quality of decisions compared to traditional methods [5,6,10]. Recently, experts have developed ML algorithms to discover hidden factors associated with HIV testing [10]. These methods are applied to any human data consisting of socio-behavioral, sociodemographic, and HIV risk information [11–13]. Machine learning can accurately predict high-HIV-risk groups that could not be identified through traditional approaches [10,14]. These techniques can enhance early prediction of HIV transmission to develop targeted testing approaches, thereby improving the efficiency of testing services and optimizing the allocation of resources [15,16]. Interestingly, ML models can also be incorporated into HIV testing devices, ultimately enhancing diagnostic accuracy [17,18].

Several ML techniques have demonstrated efficacy in precisely forecasting HIV risk and identifying the most eligible individuals for HIV testing. Least absolute shrinkage and selection operator (LASSO) models, logistic regression (LR), and random forest (RF) were used to accurately predict the HIV status of individuals based on nationwide electronic data in Demark [15]. In the United States of America (USA), ML models discovered that old age, inpatient visits, alcohol use, male gender, condom use, sexual behavior, and substance use were key predictors of HIV risk and diagnosis [19–22]. In a similar study by [23] in the USA, the incidence of HIV diagnosis among women was accurately predicted via retrospective analysis of electronic health record (EHR) data using extreme gradient boosting (XGBoost) and LASSO models. Multiple ML algorithms predicted timely clinic attendance and the uptake of HIV testing among men who have sex with men (MSM) in China [24].

Fascinatingly, ML experts are even developing novel models that can be integrated into HIV testing devices to enhance automated testing and mitigate the rate of false-positive HIV results. Demey et al. [25] applied deep learning techniques to detect 38% of false-positive HIV results in France. Additionally, in a South African pilot study, artificial intelligence (AI) conversational agents were developed on Android phones, enabling participants to conduct self-HIV counseling, risk determination, rapid testing, and linkage to care for positive cases [17]. According to the study participants, speaking with the AI agent felt normal, like talking with a human, and they felt more at ease taking an actual HIV test with the agent [17]. Beyond reasonable doubt, it is evident that ML methods significantly enhance HIV counseling and testing owing to their precise accuracy, preservation of privacy, rapidity, and facilitation of easy access to HIV testing. This review aims to understand to what extent ML algorithms can be employed to enhance the efficiency and accuracy of HIV testing interventions, as well as identify key outcomes and successes, gaps, opportunities, and limitations associated with their implementation.

2. Materials and Methods

2.1. Approach of the Systematic Review

This systematic review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA, 2020) guidelines [26]. Before starting the systematic review, the protocol was registered with the Protocol Review International Prospective Register of Systematic Reviews (PROSPERO) (ID: CRD42023464960). Thereafter, a comprehensive search was conducted for both the published and gray literature focused on utilizing ML algorithms to strengthen HIV testing globally. Published studies aimed at improving HIV testing with ML techniques were identified using developed search terms (see search strategy). A comprehensive literature search was conducted across various scientific databases, and eligible articles were screened and analyzed.

2.2. Inclusion and Exclusion Criteria

The Population, Intervention, Comparison, Outcome, and Study (PICOS) design framework [27] was used to establish criteria for selecting eligible studies for this review (the PICOS framework for this study can be found in Table S1).

2.2.1. Inclusion Criteria

This review included published scientific articles and gray literature focused on utilizing ML algorithms to strengthen HIV testing among individuals aged 15 years and above globally. In this study, we considered gray literature such as published theses, citable preprints, and book chapters. A few studies that applied predictive modeling using traditional approaches were included to compare the accuracy between ML and conventional methods. We also included studies that utilize ML in both primary and secondary data, prospectively or retrospectively predicting HIV risk, HIV status, and HIV testing, as well as studies assessing the accuracy and efficiency of ML HIV testing devices. Moreover, this review considered only studies written in English, conducted between 2010 and 2024, and available as a full text.

2.2.2. Exclusion Criteria

The review excluded published studies conducted before the year 2010, non-English studies, those not available in full text, those involving participants under 15 years of age, review articles, and studies that were not within the scope of the review.

2.3. Databases and Search Strategy

The established criteria for this review aided in developing keywords synonymous with “machine learning”, “algorithms”, and “HIV testing”. Search terms (Table 1) were generated from the Medical Subject Headings (MeSH) database and published literature on ML algorithms and HIV testing interventions. A manual literature search was conducted in different electronic databases such as PubMed, Scopus, Web of Science, Science Direct, Google Scholars, Google, and Gale OneFile: Health and Medicine. Relevant studies were further selected by searching the references of the initially identified articles. Table 1 shows search strings that were generated from electronic databases.

Table 1. Search strategy.

No	Items	Search Strategy
#1	HIV testing OR HIV diagnosis) OR HIV prevention	("HIV testing" [Mesh] OR "Testing, HIV" [Mesh]) OR "AIDS Testing" [Mesh] OR "HIV Diagnosis" [Mesh] OR "Testing HIV-AIDS" [Mesh] OR "Rapid HIV testing" OR "HIV Self-Testing" OR "HIV Prevention" OR "HIV Screening" OR "Human Immunodeficiency Virus Testing")
#2	Machine learning AND algorithms	("Machine Learning" [Mesh] OR "Artificial Intelligence" [Mesh] OR "Unsupervised Machine Learning" [Mesh] OR "Supervised Machine Learning" [Mesh] OR "Algorithms" [Mesh] OR "Models" [Mesh] OR "Deep Learning" [Mesh] OR "Neural Networks, Computer" [Mesh] OR "Support Vector Machine" [Mesh] OR "Random Forest" [Mesh] OR "Decision Trees" [Mesh] OR "Convolutional Neural Networks" OR "Recurrent Neural Networks" OR "XGBoost")
#3	#1 AND #2	("HIV testing" [Mesh] OR "Testing, HIV" [Mesh]) OR "AIDS Testing" [Mesh] OR "HIV Diagnosis" [Mesh] OR "Testing HIV-AIDS" [Mesh] OR "Rapid HIV testing" OR "HIV Self-Testing" OR "HIV Prevention" OR "HIV Screening" OR "Human Immunodeficiency Virus Testing") AND ("Machine Learning" [Mesh] OR "Artificial Intelligence" [Mesh] OR "Unsupervised Machine Learning" [Mesh] OR "Supervised Machine Learning" [Mesh] OR "Algorithms" [Mesh] OR "Models" [Mesh] OR "Deep Learning" [Mesh] OR "Neural Networks, Computer" [Mesh] OR "Support Vector Machine" [Mesh] OR "Random Forest" [Mesh] OR "Decision Trees" [Mesh] OR "Convolutional Neural Networks" OR "Recurrent Neural Networks" OR "XGBoost")

2.4. Screening Process

After searching the various databases mentioned above, the references were exported to RefWorks [28] and compiled as RIS files. The review was registered on the Covidence website where the complied RIS files were imported for screening and data extraction. The Covidence software automatically removes duplicate studies and keeps track of the entire screening process. The PICOS framework (Table S1) guided the selection of eligible articles in each stage of the screening process.

2.4.1. Phase I: Title and Abstract Screening

The first phase of the screening focused on study titles and abstracts, and only studies that applied ML algorithms to improve HIV testing interventions proceeded to the full-text screening. For studies to be eligible for full-text screening, we focused on the study population, study type, interventions, and primary objective. The authors considered study titles and abstracts demonstrating the use of ML techniques in predicting HIV risk, testing status, or the accuracy/efficacy of diagnostics tools. This was accomplished by having two independent screeners (M.J. and E.P.). For each evaluation, the two screeners independently evaluated the titles and abstracts of the qualifying articles using the inclusion and exclusion criteria, as well as a predetermined agreed-upon score criterion. In case of differences, a third member (RNPM) was asked to give input and make the final decision based on the reviews of the two members.

2.4.2. Phase II: Full-Text Screening

In the second phase, full-text screening was conducted to identify studies showcasing the efficiency/accuracy of ML algorithms in enhancing HIV testing. The full texts of the articles were read thoroughly, focusing on the study designs, methods, types of ML algorithms employed, how data were validated, evaluation techniques of models, results, and conclusions. This phase also involved two team members (M.J. and E.P.), who independently evaluated the completed the text of the qualifying articles using the criteria established by the PICOS framework (Table S1). In case of any differences, a third member (RNPM) was asked to give input and make the final decision based on the reviews of the two members. When an article was eligible for the extraction phase, a note stating the reason for inclusion was made. The same procedure applied to excluded articles.

2.5. Data Extraction and Evidence Synthesis

A data extraction tool (Table S2) was used to extract relevant information from eligible articles. Information such as study title, author names, year of publication, country, study design, study population, study sample size, outcome, type of ML algorithms used, number of models developed, details of the ML technique used to develop each model, evaluation techniques of models, key outcomes and successes, gaps, opportunities, and limitations associated with ML implementation, and key findings are shown in tables. This process was guided by the same PICOS (Table S1) framework used to define the eligibility criteria of this study on the Covidence web software. The extracted information was presented in tables and graphs, and the results were narratively summarized.

2.6. Quality Appraisal

Quality appraisal was conducted using the Critical Appraisal Skills Program Systematic Review (CASPSR) checklist (Table S3), which allows for the investigation of a study's validity, precision, and generalizability. Two reviewers (M.J. and E.P.) screened all eligible studies independently. They evaluated the complete text and abstract of the qualifying articles using the inclusion and exclusion criteria and a predetermined and agreed-upon score criterion for each evaluation. In case of any differences, a third member (RNPM) was asked to give input and make the final decision based on the reviews of the two members. The PRISMA checklist was also used to ensure that all the components of the systematic review were transparent and documented (Table S4).

2.7. Analysis

The shreds of evidence extracted from the eligible studies were summarized in figures and tables, and then key themes were used to describe the important findings. The first part of the results section consists of the studies' characteristics, such as authors' information, geographical distributions of studies, publication trends, study types, and populations. The subsequent sections narratively summarized the types of ML algorithms, their performances, key findings, strengths, limitations, and future directions of ML models employed in HIV testing interventions. We compared the accuracies of traditional methods (TMs) and ML models. Empirical differences between the studies that applied both TMs and ML within our sample were evaluated. An independent t-test was used to determine if there was a statistically significant difference between the accuracies of ML and TM, where the significance level was set at $p < 0.05$ using 95% confidence intervals (CI). Statistical analysis was performed using STATA version 18.

2.8. Summary of the Systematic Review Process Using the PRISMA Guidelines

In this review, a total of 2261 articles were found through the initial search conducted on different databases (PubMed = 804, Web of Science = 456, Google Scholar = 444, Scopus = 345, Science Direct = 198, Gale OneFile = 6), Google = (3), and Citation search (=5). After removing duplicates, 845 articles remained, which went through title/abstract screening, and ineligible articles were also removed. Thereafter, 218 eligible articles were further screened for full-text review, out of which 51 articles were retained for the final systematic review. The study selection process is shown in Figure 1.

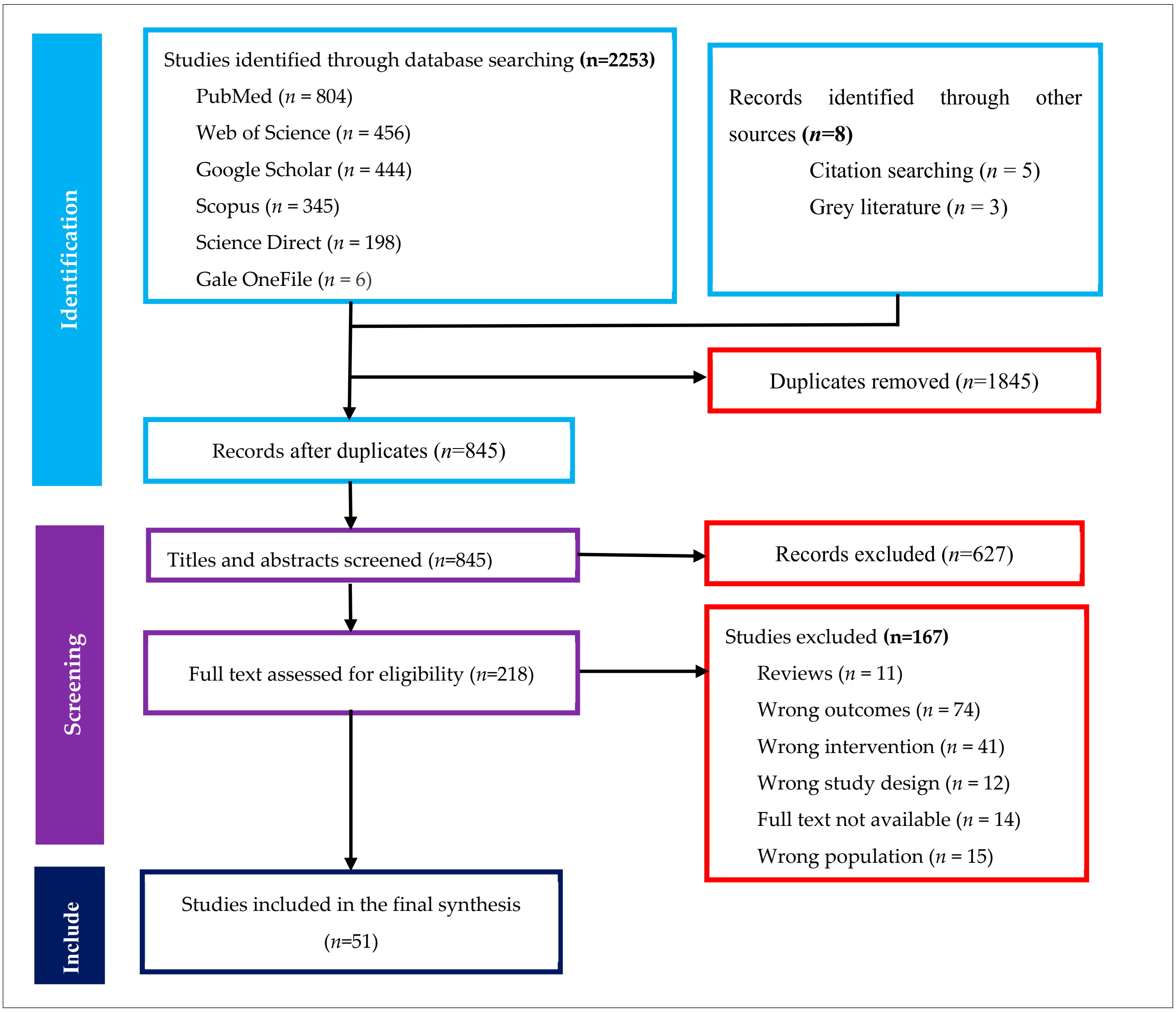


Figure 1. PRISMA flow chart representing the review selection process. Note: Reviews—any form of review article; Wrong intervention/outcome—uses of ML in CD4 and viral load testing, uses of ML in other HIV prevention and treatment protocols other than HIV testing, papers without results, and papers that reported findings related to ML in relation to HIV risk, status, and testing; Wrong design—studies without an ML approach or predictive modeling comparable to that in ML; Wrong population—studies on individuals less than 15 years old.

3. Results

3.1. Characteristics of the Selected Studies

This review analyzed 51 eligible studies, which are narratively summarized below. The first section of the narrative analysis describes the study characteristics, focusing on author information, geographical region, year of publication, study types and designs, study population, and sample sizes (see Table S5: a tabular summary of the study characteristics).

3.1.1. Geographical Distributions of the Selected Studies

The selected 51 studies were conducted across various regions of the world. Figure 2 shows the continental distributions of the selected studies, and Figure 3 shows countries with the highest number of studies applying ML to strengthen HIV testing within our sample. Most of these studies were produced in the Americas (*n* = 21), followed by Africa (*n*=18) and Asia (*n* = 8) (Figure 2). Comparing the number of studies by country, the USA (n = 18) has the highest, followed by China (*n* = 5), Zimbabwe (*n* = 5), Ethiopia (*n* = 5), Australia (*n* = 4), and South Africa (*n* = 3) (Figure 3). The list of all the countries and their number of studies can be found in Table S5.

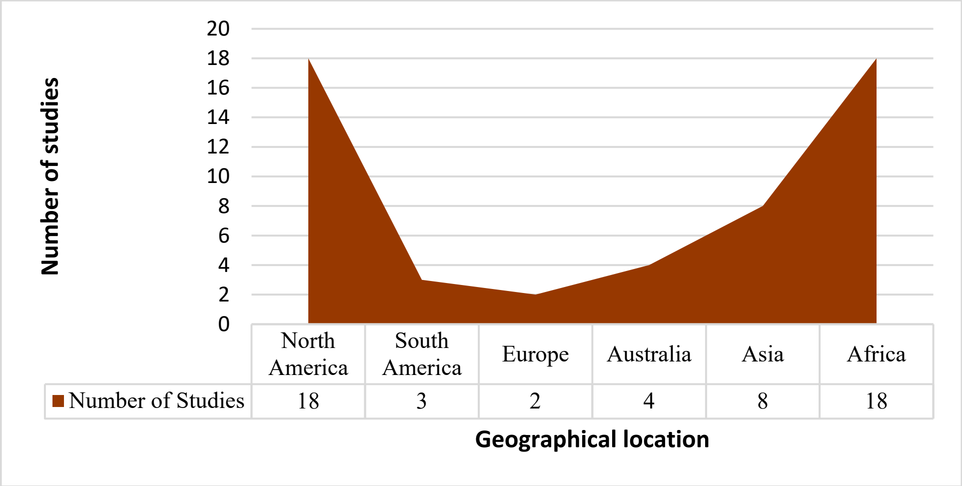
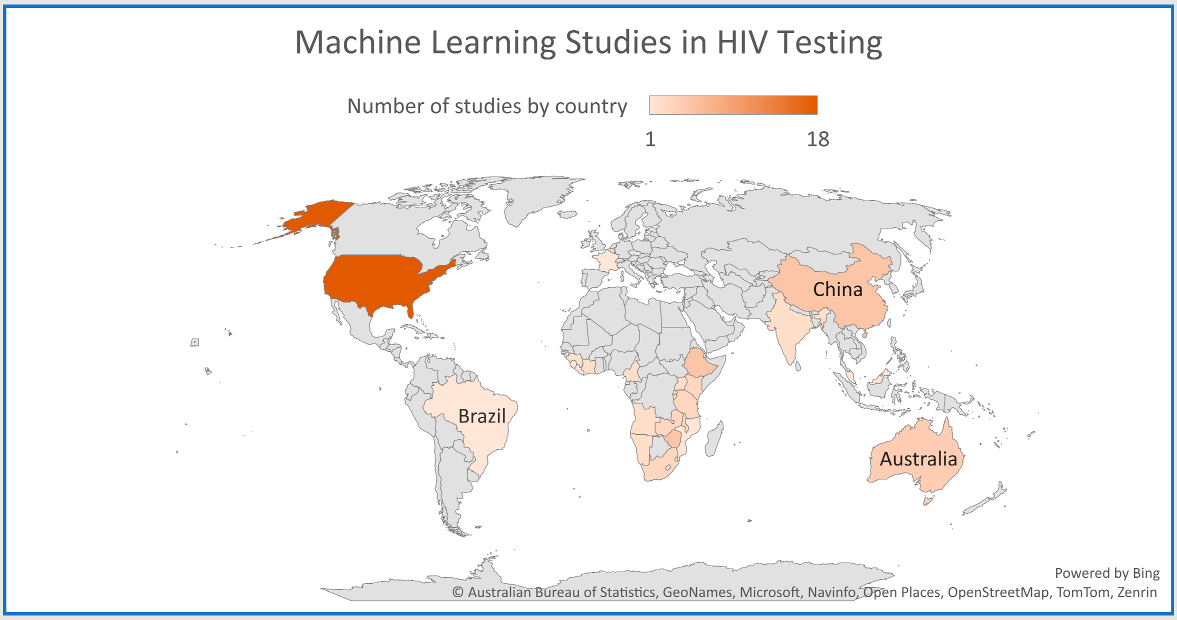


Figure 2. Distribution of studies by region. Note: Seven studies were conducted in multiple countries.



Top 10 countries with a minimum number of 3 studies

No.	Country	Number of Studies	No.	Country	Number of Studies
1	USA	18	6	South Africa	3
2	China	5	7	Tanzania	3
3	Zimbabwe	5	8	Kenya	3
4	Ethiopia	5	9	Zambia	3
5	Australia	4	10	Malawi	3

Figure 3. Distribution of studies by country.

3.1.2. Yearly Publication of the Selected Studies 2010–2024

This review analyzed 51 eligible articles that were published between 2010 and 2024. Figure 4 exhibits increased trends of studies applying ML in HIV testing interventions between 2010 and 2024. The highest number of publications was between 2020 and 2024 (74.5%; *n* = 38), while the lowest trends were observed between 2010 and 2014 (3.9%; *n* = 2) (Figure 4).

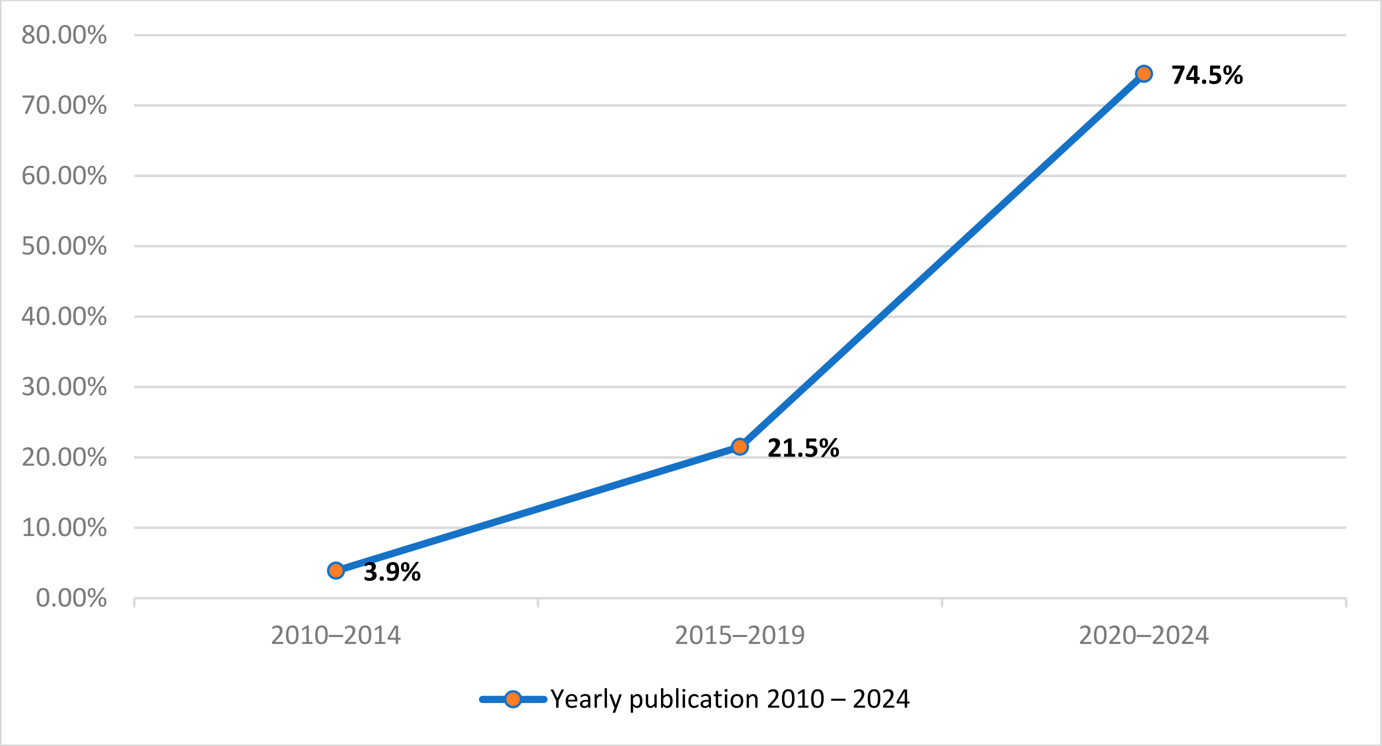


Figure 4. Yearly publication of the selected studies.

3.1.3. Additional Characteristics of the Selected Studies

Figure 5 shows that the selected studies comprised both general and KPs (i.e., PLHIV and MSM) aged 15 years and above. However, a study involving individuals aged between 12 and 80 years was included in the analysis. Although the lower age group falls below the threshold of 15 years of age outlined in our inclusion criteria, the study's comprehensive content and alignment with our research objectives warranted its inclusion. Furthermore, 36 studies were conducted among general and PLHIV, and 15 involved MSM. Most (37) studies conducted cross-sectional retrospective analyses using secondary data. Only a few (14) were primary studies, with 10 being experimental/intervention studies, 1 being a nested case–control study, and 3 being cross-sectional. The most common secondary data types included demographic and health surveys (DHS), EHR, and social media data. The smallest sample size was 10 participants, while the largest was 4,384,178 participants (Figure 5). The complete list of characteristics of studies included in the analysis can be found in the Supplementary Materials (Table S5).

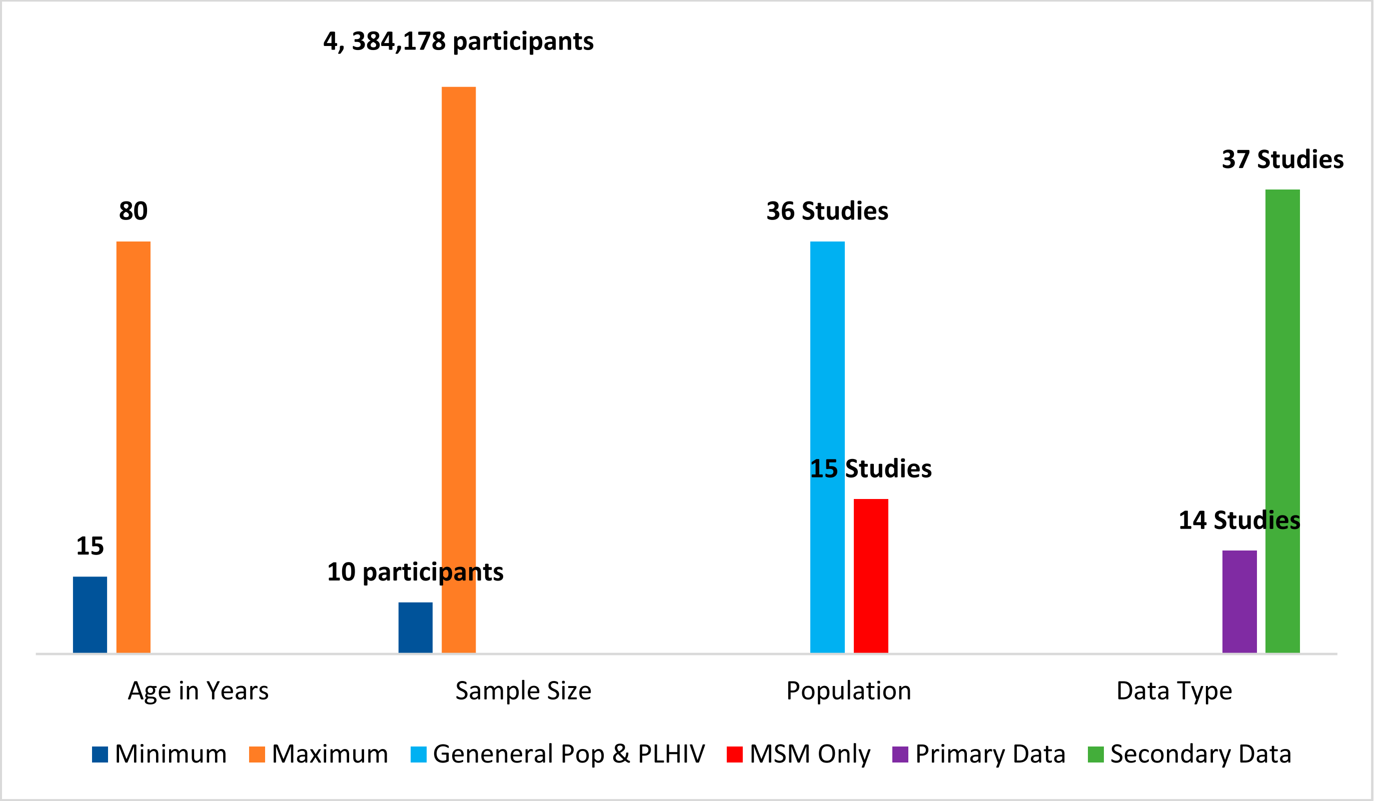


Figure 5. Study characteristics. PLHIV: people living with HIV; MSM: men who have sex with men.

3.2. Machine Learning Interventions for Enhanced HIV Testing

3.2.1. The Machine Learning Algorithms Applied

The studies included in this review employed various ML algorithms and interventions to enhance HIV testing, along with key findings from the selected studies. Of the 51 articles, 47 utilized several ML types, including supervised and unsupervised learning, and some studies applied both TMs and ML models. Four studies that developed predictive models using only traditional statistical methods were included in the sample. Across all the chosen studies, more than 30 distinct commonly used ML algorithms were implemented. A few studies developed novel ML models such as HealthPulse AI, MySTIRisk, and SexPro for effective HIV risk assessment and improved testing.

Most Frequently Utilized Machine Learning Algorithms in HIV Testing Interventions

The results (Table 2) reveal that the most frequently used ML algorithm was LR, a supervised machine learning (SML) technique applicable for classifications [24], featured in 22 studies with high accuracy [14,15,19–21,24,29–44]. Random forest, another SML classifier [24], was the second most common algorithm, exhibiting exceptional performance in nearly all the 17 studies in which it was applied [10,13,15,19–21,24,29,31,32,34–36,41,42,45,46]. Support vector machine (SVM), an SML algorithm, ranked third and was employed in 14 studies [10,20,22,24,30–32,36,39–42,47,48].

Table 2. Most frequently utilized machine learning algorithms in HIV testing interventions.

ML Algorithms	**No. of Studies**	**Largest Dataset Size/Reference**	**Best Performance/Reference**	**Binned Size**	**References**
LR	19	4,348,178 [15]	90.5% highest accuracy Best model in [14,36]	Large	[14,15,19–21,24,29–44]
RF	15	4,348,178 [15]	94.4% highest accuracy Best model in [13,19,32,34,36]	Large	[10,13,15,19–21,24,29,31,32,34–36,41,42,45,46]
SVM	13	124,777 [40]	95.1% highest accuracy Best model in [22]	Large	[10,20,22,24,30–32,36,39–42,47,48]
XGBoost	9	124,777 [40]	99% highest sensitivity Best model in [10,23,24,35,40]	Large	[10,13,23,24,29,30,34,35,40]
DL models (ANN, RNN, CNN, LSTM)	9	88,642 [13]	98% highest accuracy Best model in [47–49]	Large	[13,24,30,31,47–51]
LASSO model	7	4,348,178 [15]	82% highest accuracy	Large	[15,22–24,37,52,53]
DT	7	56,682 [25]	81.9% highest accuracy Best model in [31]	Large	[14,25,31,32,35,39,50]
KNN	6	6672 [35]	80% highest accuracy	Small	[10,24,29–31,35]

Note: LR, logistic regression; RF, random forest; SVM, support vector machine; XGBoost, extreme gradient boosting; DL, deep learning; ANN, artificial neural network; RNN, recurrent neural network; CNN, convolutional neural network; LSTM, long–short-term memory; LASSO, least absolute shrinkage and selection operator; DT, decision tree; KNN, K-nearest neighbor.

Furthermore, XGBoost was ranked the fourth most common algorithm and was employed in nine studies [10,13,23,24,29,30,34,35,40]. Moreover, nine studies employed various deep learning (DL) models, including the artificial neural network (ANN), recurrent neural network (RNN), convolutional neural network (CNN), and long–short-term memory (LSTM) [13,24,30,31,48–51]. The sixth most common algorithm was LASSO, a regression model that combines variable selection and regularization to improve prediction accuracy and interpretability [54]. The LASSO model and decision tree (DT) were each implemented in seven studies, and the K-nearest neighbor (KNN) algorithm appeared in six studies [10,14,15,22–25,29–32,35,37,39,41,52,53].

The least utilized models, including both TMs and ML algorithms, were as follows: gradient boosting machines (GBMs), light gradient boosting machines (LGBMs), elastic net (EN), naïve Bayes (NB), principal component analysis (PCA), HealthPulse AI, adapting boosting (Adaboost), extra tree, partial least squares discriminant analysis (PLS-DA), class analogy (SIMCA), MySTIRisk, SexPro, bagging classifier (BC), linear discriminant analysis (LDA), generalized additive model (GAM), classification and regression tree (CART), AI conversational agent, baseline model, unigram model, graph convolutional network (GCN), generalized linear models (GLMs), nomogram, AI Chatbot, association rule algorithm, greedy algorithm, multivariate ML, AI algorithm, Cox proportional hazard regression model, and extreme-gradient-boosted tree model (xgbTree) [10,15,17,20–22,24,29–31,34,35,37,40,42,46,50,55–67]. The list of all the ML algorithms in the various studies analyzed in this review can be found in the Supplementary Files (Table S6).

Best Performing Machine Learning Algorithms

Several studies utilized multiple algorithms, and their performance metrics are reported in Table 2. XGBoost demonstrated the highest performance compared to other algorithms in [10,23,24,35,40]. Additionally, RF emerged as the best-performing algorithm compared to other models in the reviewed studies [13,19,32,36]. The DL models, like RNN and CNN models, achieved the best performances in several studies [47–49]. The remaining algorithms that achieved the best performance metrics included DT, GBM, SVM, LR, GCN, nomogram, and xgbTree [14,21,31,34,36,37,42].

3.2.2. Machine Learning Interventions Employed to Enhance HIV Testing

In this review, studies were given priority based on their focus on ML interventions within specific thematic areas. These thematic areas include the prediction of HIV risk/test/status, evaluation of the accuracy and efficiency of ML in HIV testing compared to TMs, and novel ML techniques for enhanced HIV testing uptake, including HIVST. Furthermore, several studies investigated and identified predictors related to HIV risk, status, and testing. The ML interventions are elaborated in the following thematic contexts:

- Predicting HIV risk/testing: As summarized in Table 3, numerous studies predicted individuals at the highest risk of HIV, as well as those who were likely to undergo HIV testing. This study's central focus on predictive modeling within the context of HIV testing has unearthed compelling evidence showcasing the effectiveness of ML in enhancing HIV testing interventions. For instance, [15] utilized LASSO models, LR, and RF to predict the HIV status of individuals based on nationwide electronic data. The findings demonstrated enhanced predictive performance, estimating that 384 individuals would need to be tested to identify one undiagnosed person with HIV [15]. A parallel study by [51] used DL models to forecast HIV incidence between 2022 and 2023 in the Philippines. The study predicted a cumulative case count of 145,273 by 2030 [51].

Table 3. Machine learning interventions employed to enhance HIV testing.

Machine Learning Interventions	Key Findings	References
• Predicting HIV risk/testing • Socio-behavioral and demographic predictors of HIV	Models were developed; ML models that accurately predicted individuals at the highest risk of HIV for prioritized testing were validated -HIV diagnosis was accurately predicted from secondary data A study predicted monthly incidences of HIV from 2020 to 2030 -Biomarker indicators for HIV diagnosis were identified Dynamic changes in the immune state were associated with HIV Age, gender, race, wealth index, province, contraceptive use, sexual behavior, condom use, alcohol/drug usage, knowledge of HIV, intimate partner violence, hepatitis diagnosis, pregnancy, depression, male circumcision, being male, MSM, and STI history/diagnosis as key predictors of HIV Studies developed models such as MySTIRisk and SexPro for self-HIV risk assessment -The number of those who needed to be tested to find one undiagnosed PLHIV from certain communities/countries was predicted HIV diagnosis from magnetic resonance imaging (MRI) was predicted using ML -HIV prevalence and incidence was forecasted	[10,13,15,19–22,24,30–33,35–38,40,43–48,50–53,56,58,59,61,62,65–68]
• Early HIV diagnosis	Models that enhance early HIV diagnosis using multiple ML from secondary data were developed Models predicted the number of people to be diagnosed with HIV in the next decade by analyzing their socio-behavioral data Individuals' HIV statuses were predicted using DHS, EHRs, and hospital records HIV hotspots in multiple countries were identified ML models improved workflow with the ability to report immediately to reduce infection	[22–25,29,49,53,68]
• Accurate/Efficient HIV testing • Cost and resource optimization in HIV testing • Misinterpretation of HIV test results	ML is more likely to correctly interpret true-negative HIV test results (NPV = 100%) ML models improve diagnostics performance and reduce false-positive/-negative results HealthPulse AI provides accurate and consistent results in interpreting HIV RDT test results ML models that are efficient and effective for HIV testing among KPs such as MSM were developed ML methods outperformed traditional approaches in several studies Studies developed ML algorithms that can accurately facilitate HIV counseling and testing as well as interpret HIV test results with and without internet connectivity Cost-effective HIV testing for low-income countries was achieved An ML method increased the economic benefits of HIVST kit distribution by more than 23%	[22,29,33,34,48–50,57,59,63]
• Enhancing HIV-Self Testing • Innovative approaches in HIV diagnostic tools	ML models enhance SHIVT and improve interpretations of results Chatbots and conversational agents with mHealth solutions improve convenience for using HIVST Models diagnosed HIV from MRI using ML	[17,29,41,60,64,65]

Note: HIV, human immunodeficiency virus; HIVST, HIV self-testing; ML, machine learning; MRI, magnetic resonance imaging; DHS, demographic and health survey; EHRs, electronic health records; MSM, men who have sex with men; STI, sexually transmitted infections.

A nested case–control study in Peru utilized LASSO ML to predict HIV acquisition, incorporating selected biomarkers and socio-behavioral covariates among MSM and transgender women [52]. The findings underscored that changes in immune states are associated with HIV acquisition, and the inclusion of biomarker, demographic, and behavioral data provided complementary information to assess HIV risk [52]. Similarly, several SML algorithms were applied to the Zimbabwean DHS dataset to predict HIV positivity [35]. The study highlighted that condomless sex, multiple sex partners, a lack of understanding about HIV transmission, being unemployed, and intimate partner abuse were all related to HIV transmission [35]. From a retrospective analysis of EHRs, the XGBoost and LASSO models were able to accurately predict the incidences of HIV diagnoses among women in the USA [23]. Various studies have also used EHRs, DHS, and clinical data to predict HIV risk/testing based on sociodemographic and socioeconomic variables in both general and KPs [10,13,14,20–22,32,36–38,40,47,48,50,53,58,59].

Machine learning can predict those at higher risk of HIV infection, improve optimal HIV screening choices, and facilitate HIV testing [69]. Our analysis identified sociodemographic, socio-behavioral, and other factors associated with increased HIV risk for prioritized testing. A study by the authors of [52] discovered that dynamic changes in immune states are associated with HIV acquisition, and biomarker, demographic, and behavioral data added complementary information to HIV risk. Studies from the USA highlight factors such as old age, inpatient visits, alcohol use, male gender, condom use, sexual behavior, and substance use as key predictors of HIV risk and diagnosis [19–22]. Furthermore, studies on MSM in China and Australia used different algorithms to predict HIV risk and timely HIV testing [24,32]. Triage reasons such as STI symptoms, asymptomatic screening, sex overseas in the previous 12 months, PLHIV, contact with someone with the infection, current sex work, screening reminder frequency, type of screening reminder method, sex with a male in the previous three months, and condom use were among the major predictors [24,32].

Most of the studies from SSA were conducted in East and Southern Africa, including the Population-Based HIV Impact Assessment (PHIA) countries. The findings of [10,14,31,40,42,58] showed that age, relationship with the family head, the highest level of education, the highest grade at that school level, work for payment, avoiding pregnancy, age at the first experience of sex, wealth quintile, male circumcision, sex, province, sexual debut, sexually transmitted infection (STI), race, wealth index, contraceptive use, longitude, latitude, altitude, knowledge of HIV, place of residence, age of the most recent partner, total lifetime number of sexual partners, years lived in the current place of residence, having receptive vaginal sex, and condom use during the last intercourse were significant predictors of HIV status.

In the In the Mano River Union countries (Liberia, Sierra Leone, Cote d’Ivoire, and Guinea), Cote d’Ivoire was reported as the hotspot for HIV prevalence between 2005 and 2015, while Guinea and Sierra Leone became hotspots for HIV prevalence from 2016 to 2020 [53]. In addition, the study indicated that HIV prevalence was higher among women than among men who lived in urban areas. The model used in the study further predicted that comprehensive correct knowledge about HIV was below the UNAIDS threshold for ending HIV endemic [53].

- Accurate and efficient HIV testing interventions utilizing machine learning techniques: This review encompasses studies that employed novel technologies and algorithms to enhance HIV testing accuracy and efficiency. In a study conducted in France, Demey et al. [25] employed DL techniques to enhance the efficiency of HIV diagnostic tests. The research involved a retrospective analysis of “Centaur$^{®}$ CHIV” assays and confirmatory tests conducted at Amiens University Medical Center between 2012 and 2018 [25]. The findings revealed the detection of 38% false positives and 62% confirmed true positives [25]. Remarkably, the models demonstrated high accuracy by significantly reducing the number of false-positive CHIV assay results from 171 to 12 [25]. Similar studies conducted in the USA developed models that effectively reduced HIV false-positive results [22,57]. One of the South African studies we reviewed developed AI conversational agents for automated HIV self-counseling and testing [17]. The AI conversational agents/models were developed on Android phones, enabling participants to conduct self-HIV counseling, risk determination, rapid testing, and linkage to care for positive cases. Out of the ten participants in the study, six testers found that talking to the agent felt natural and equivalent to chatting with a human. At the same time, seven said they would feel comfortable taking a real HIV test with the agent. The study argues that this method is more effective than traditional HIV testing and counseling due to its anonymity, privacy, speed, and easy access [17].

The findings of this study underscore a significant demand for the application of mobile-based algorithms for efficient, accurate, automated, and cost-effective HIV testing. An inbuilt mHealth model, combined with DL, demonstrated the ability to reduce false-negative HIV RDT test results with high levels of sensitivity and specificity compared to traditional HIV rapid diagnostic tests in South Africa [48]. HealthPulse, another AI-based mobile algorithm designed to detect and interpret RDTs for HIV in low- and middle-income countries (LMICs), not only enhances RDT efficiency but also prioritizes privacy, accountability, and accessibility [50,55]. Moreover, ML interventions have demonstrated high accuracy in HIV prediction and detection in Ethiopia, Kenya, and Uganda [33,49,55] (Table 3; also see Table S6).

- Enhancing the uptake of HIV self-testing among MSM using machine learning: The WHO recommends HIVST as a convenient and confidential option for HIV testing, especially among KPs who are often reluctant to access facilities for counseling and testing services [7]. This review synthesized studies that used ML to improve the uptake and utilization of HIVST among MSM from different countries. In China, an ML approach increased the efficiency of HIVST by 17.7%, and the distribution of HIVST kits was improved by 18% [39]. In a mixed study from Malaysia, an AI Chatbot was compared with human HIVST counseling and testing [60]. The study revealed that 93% of the participants perceived that the AI Chatbot provided more comprehensive information about testing, 100% of them stated that the chatbot was more convenient, and 79% were willing to continue using it [60]. Jing et al. [63] used a Greedy algorithm to increase the economic benefits of the secondary distribution of HIVST kits by more than 23% compared to those achieved by conventional methods. However, the uptake of HIVST is still low in many developing countries as the services are more friendly to individuals who can read and comprehend English. One of the recommendations from Cheah et al. [60] is to develop ML-aided HIVST technologies suitable for illiterate people. This study agrees with these recommendations since the novel technique is required more in developing countries like SSA, with its high HIV prevalence and high rates of illiteracy among its population.

3.3. Comparison between Traditional and Machine Learning Predictive Modeling in HIV Testing

3.3.1. Empirical Difference between Traditional and Machine Learning Methods

Table 4 compares nine studies that applied both TMs and ML for enhancing HIV testing. Machine learning outperformed both human and traditional statistical methods employed in seven out of the eight studies. The results highlighted that ML models are more precise and accurate in HIV testing predictive modeling than traditional LR [32,39]. Moreover, a study by [55] reveals that ML methods are more effective when interpreting true-negative test results than humans, while humans exhibit higher performance in interpreting true-positive test results. However, there was a small overlap between the positive predictive values (PPV) and negative predictive values (NPPV) of the two methods. Furthermore, ML approaches consistently outperformed TMs in enhancing the uptake of HIVST among MSM [39,41,63]. Even though a limited number of studies were used for this comparison (Table 4), our findings are consistent with empirical evidence from other studies suggesting that ML provides better predictions.

Table 4. Empirical differences between machine learning and traditional methods in HIV testing predictive modeling.

Studies	Traditional Methods	Machine Learning	Outcome
Roche et al. [55]	Human interpretation of HIVST	AI algorithm interpretation of HIVST	Humans are more likely to correctly interpret true-positive HIV test results (PPV = 100%) compared to the AI algorithm (PPV = 82%) The AI algorithm (NPV-100%) is more likely to correctly interpret true negatives in comparison to humans (NPV-99.9%)
Ni et al. [41]	Human interpretation and empirical models	ML models	The ML model motivated more individuals to conduct HIVST The difference between the ML model and the empirical scale was not significant
Jing et al. [39]	Human	ML models	The ML model outperformed human identifications and distribution of HIVST kits The ML approach increased HIVST kit distribution by 18%
He et al. [32]	TMs	ML models	-The ML models outperformed TMs in HIV risk prediction among MSM -ML achieved 94% accuracy
Jing et al. [63]	TMs	ML	The ML method outperformed TMs The ML method increased the economic benefits of HIVST kit distribution by more than 23%
Bao et al. [34]	TMs	ML	The ML models consistently outperformed the TM used The ML achieved an accuracy of 76.3% compared to that of TMs (68%)
Oladokun et al. [14]	TMs	ML	The TM used outperformed the ML model Both models did not achieve high accuracy
Rice et al. [66]	TMs	ML	The ML model outperformed the TM used HIV testing was increased by 18.8% in the AI group compared to the comparison group (8.1%)
Balzer et al. [33]	TMs	ML	ML outperformed TMs ML model achieved an accuracy of 78% compared to TM (68%)
Conclusion	TMs outperformed ML in only two studies	ML outperformed TMs in seven studies	ML models outperformed TMs based on the empirical evidence from the study's sample

Note: AI, artificial intelligence; HIV, human immunodeficiency virus; HIVST, HIV self-testing; ML, machine learning; MSM, men who have sex with men; TM, traditional methods; PPV, positive predictive values; NPPV, negative predictive values.

3.3.2. Statistical Differences between Traditional and Machine Learning Models

A total of 43 studies involving the use of either ML or TMs as predictive models for HIV risk/testing were sampled. The results of the independent sample *t*-test indicated a statistically significant difference between the accuracy of ML and TMs (Table 5). On average, the sampled studies from different parts of the world that utilized ML reported a 12.53% ($p = 0.0002$) higher accuracy than selected studies involving the use of traditional predictive models. Therefore, our findings indicate that ML models are more accurate than TMs in predicting HIV risk/testing.

Table 5. *t*-test results of the comparison between traditional and machine learning.

Predictive Model	Studies	Mean	Std. Err	Std. Dev	95% CI
Machine Learning	35	86.66	1.79	10.61	82.01–89.03
Traditional Method	8	73.13	2.29	6.47	67.72–78.54
Combined Values	43	83.33	1.69	11.07	79.92–87.73
Difference in Values		12.53	2.91		6.39–18.66

Std. Err: standard error; Std. Dev: standard deviation; 95% CI: 95% confidence interval.

3.4. Successes of Machine Learning Interventions in HIV Testing

This study reveals that ML has shown promising success and opportunities in HIV testing interventions (Table 6). The results showed that the employed ML algorithms demonstrated remarkable performance in precisely identifying the most eligible individuals for HIV testing. The ML models enhance diagnostic tools and provide cost-effective options for HIV testing across different regions. In comparison to TMs, ML can analyze complex datasets more rapidly and accurately [69,70]. Additionally, our study affirms that ML models can successfully reduce false negatives and false positives in HIV testing, thereby increasing the reliability of test results [22,25,48,56].

Table 6. Successes, gaps, and future directions of machine learning interventions for enhanced HIV testing.

Features	Successes/Strengths	Gaps/Limitations	Future Research and Development
Study design	14 studies utilized primary data, and most of them prospectively predicted HIV testing. This is essential for the following: Minimizing missing information Preventing recall bias Capturing important sociodemographic and socio-behavioral data	Most of the studies employed cross-sectional retrospective designs using secondary data, which are subject to the following: Recall bias Incomplete information Imbalanced data	More studies should evaluate the effectiveness of ML models by using prospective designs Primary studies are essential for capturing real-time information and ensuring high data quality
Data quality and sources	Some studies used large samples, providing enough data to train the ML models Studies that collected primary data were more likely to have quality data to provide accurate predictions	EHRs and hospital records sometimes have limited patient information Many studies used cross-sectional survey data Self-reported information is subject to recall bias Secondary data, in general, are subject to missing information Many studies experienced data imbalances	Data used for HIV testing predictive modeling should be cleaned Missing data from datasets should be properly treated The use of primary data should be encouraged for this kind of predictive modeling
ML models	Studies developed and validated models that are highly accurate in predicting HIV testing Studies developed novel models for the first time with high accuracy	Some ML models were overfitted Some models are complex to interpret Poorly developed models can lead to false prediction	New models need to be trained and validated with different sets of data Training on different sets of models, especially in developing countries, needs to be improved
Generalizability	Studies evaluated the accuracy of several MLs on a single dataset Studies utilized very large sample sizes Studies used datasets from multiple countries to predict HIV testing	Many studies applied a single ML algorithm Some samples were very small for predictive modeling Many studies were conducted on a single population/country	ML models should be trained using large amounts of data with a variety of variables ML models with high accuracy should be evaluated using multiple-country data
Ethical Considerations of ML	Some ML models analyzed large amounts of data while addressing data security and privacy concerns ML enhances HIVST for individuals with privacy concerns	ML, in general, raises ethical concerns, and many studies fail to address this There is a lack of trust in and acceptability of ML methods Some ML models still require improvement to facilitate self-counseling and effective testing for HIV	ML models should be improved, ensuring high data protection and privacy-conscious ML applications used for HIVST should be improved to avoid ethical concerns with HIV testing
Accuracy and performance	ML models are highly accurate and outperformed TMs in several studies ML models accurately classify false-positive/-negative results	-Some ML models were less accurate than traditional predictions Some ML models are only effective in predicting either false positives or false negatives	ML models require more improvements with adequate data to achieve consistent accuracy beyond human capacity

Note: EHRs, electronic health records; HIV, human immunodeficiency virus; HIVST, HIV self-testing; ML, machine learning; TM, traditional methods.

Another advantage of ML interventions highlighted in this study is their ability to forecast future trends and patterns of HIV/AIDS, which enables the early detection of HIV, allowing for timely intervention and improved management of the infection [13,33,38,59]. Moreover, ML algorithms can adapt and customize testing approaches based on individual risk factors, optimizing the testing process for different populations (including KPs) and improving the overall effectiveness of HIV testing programs [13–15,19–21,24,29–32,35–40,47–53,56,59,68].

Fascinatingly, ML algorithms can be integrated into mobile-based applications, providing opportunities for widespread and accessible HIV testing. This is very beneficial in remote or underserved areas where access to traditional healthcare facilities may be limited. It also improves HIVST automation to boost privacy, confidentiality, accessibility, and acceptance of HIV testing. This is particularly important in the context of HIV testing, where stigma and discrimination pose significant obstacles to the KPs [17,48,50,55]. This study highlights that ML has the potential to be integrated into a variety of health technologies, such as telemedicine and electronic health records, resulting in a comprehensive and interconnected healthcare system. This integration has the potential to improve healthcare professionals' collaboration and communication, resulting in a complete approach to HIV testing and care [23,38,59,68].

3.5. Gaps Identified via the Application of Machine Learning Interventions in HIV Testing

While ML has demonstrated potential in the realm of HIV testing interventions, numerous challenges, limitations, and gaps warrant careful consideration. This study summarizes the gaps identified in Table 6. The efficacy of ML models heavily relies on the quality and representativeness of the training data. Biases within the training data, especially when they lack diversity or representativeness, can result in skewed models, potentially leading to inaccurate predictions and disparities in HIV testing outcomes (see Table S7).

Moreover, a notable limitation identified from the reviewed studies is the restricted generalization of models. As such, models employed in a specific region or on a particular demographic may not be transferable to other regions or populations. Most studies examined in this review utilized retrospective designs, incorporating self-reported information vulnerable to missing data and recall bias. While some models experienced overfitting, a few of them were underfitted; in either case, the accuracy of the results is compromised. Additionally, several studies had from small sample sizes, lacking representativeness [14,17,41].

Despite the various studies advocating for the integration of ML interventions in healthcare systems, obstacles such as compatibility issues, insufficient expertise, and the necessity for specialized training for healthcare professionals, particularly in low-resource settings, pose impediments to seamless integration. For example, studies have developed ML models based on mobile platforms to enhance HIVST and automate the interpretation of HIV results. However, these models may face usability challenges in low-resource settings due to insufficient expertise [17,20,48,60].

Additionally, some ML models, especially DL models, are often considered "black boxes" due to their complexity, which can be challenging to interpret. This makes it difficult to understand how they reach a particular prediction [70]. The intricacy of ML models and their unclear predictive capabilities raise ethical concerns surrounding data privacy. The vast majority of ML models predicting HIV rely on secondary data, such as hospital records. However, there is limited evidence of robust defense mechanisms in AI/ML systems to safeguard integrity, leading clinicians to question the reliability of integrating ML techniques into healthcare [70,71].

Furthermore, this review identified a significant discrepancy in the application of ML interventions in the context of HIV testing in SSA and developed nations. Of the 51 studies included in this analysis, 18 were conducted in SSA, of which 16 were from East and South Africa, with only two studies involving West and Central African countries. This necessitates the promotion of ML integration into HIV testing interventions across other African regions for improved HIV testing (Tables 6 and S7).

3.6. Future Directions with the Utilization of Machine Learning for Enhanced HIV Testing

This review suggests that ML interventions are essential for achieving the UNAIDS 95-95-95 goal of eradicating AIDS by 2030, as they can accurately predict high-risk populations and efficiently target cost-effective interventions. Despite the various challenges associated with their implementation, studies have made valuable recommendations for improving the acceptability and usability of ML interventions, especially in low-resource settings.

Some studies have developed novel models for the first time, believing that if this approach is replicated by authors of other studies in different settings, it would increase the generalizability of the models. Furthermore, studies propose the application of ML algorithms in prospective studies to yield more meaningful data that can accurately predict individual behavioral risk without any recall bias.

Although ethical concerns with the applicability of ML in healthcare persist, emerging studies have devised models aimed at improving data privacy [17,46,50,70,71]. Some of these studies emphasized the importance of prioritizing data privacy, and proposed that such models should be replicated in healthcare settings [14,17,38,43,46,50,58,59,70,71].

It is also important to mention the need for more professional training in ML for improved usability and acceptability. Improving the accuracy of ML models is another important subject. He et al. [32] emphasized the need for the regularization and minimization of model complexity to address overfitting. Moreover, more studies involving ML in HIV testing interventions should be conducted across SSA, especially in high-HIV-incidence epidemic settings [32].

4. Discussion

4.1. Main Findings

This study examines different ML algorithms that have been employed in various studies across the globe aimed at enhancing the efficiency and accuracy of HIV testing interventions, and identifies key outcomes, successes, gaps, opportunities, and limitations associated with their implementation. This review included 51 studies, of which 21 were conducted in the Americas, 2 were from European countries, 4 were from Australia, 9 were from Asia, and 18 were from SSA. In terms of individual countries, the USA contributed the highest number of studies (n = 18), followed by China (n = 5), Zimbabwe (n = 5), Ethiopia (n = 5), Australia (n = 4), and South Africa (n = 3). Among the 18 studies conducted in SSA, 16 were in East and South Africa, while only 2 [46,53] were conducted in West and Central African countries. Several factors could explain these disparities. The USA is one of the most developed countries globally, having achieved significant advancements in research and technology [72]. In addition, East and South Africa bear the highest global HIV burden [4], possibly explaining the concentration of HIV-related studies in these regions.

Most studies utilized multiple SML models, with LR, RF, SVM, the LASSO model, XGBoost, DT, and DL being the most utilized algorithms, except for a few studies that employed UML. In contrast to UML, SML allows the researcher to easily understand the actual classes in the training data even prior to inputting the data for training [73]. Despite the challenges in classifying big data, SML provides more accurate results, making it the most popular ML approach used in various fields of research [74].

The best-performing algorithms were XGBoost, RF, DL (RNN and MobileNet2-CNN), DT, GBM, SVM, LR, GCN, and monogram. In most studies, these models outperformed other models in terms of accuracy. The XGBoost model outperformed several algorithms, exhibiting high predictive performance toward strengthening HIV testing interventions. However, our findings contrast those of Ferreira et al., who examined supervised automated machine learning (SAML), DL, and XGBoost and discovered that SAML had the most significant predictive performance [75]. The datasets used in their investigation were relatively small, which may explain the inconsistencies [75]. Furthermore, this review found that RF outperformed other models in predicting HIV and HIV testing. Random forest outscored SVM, XGBoost, LR, DT, and traditional approaches in various studies [13,19,36]. This is supported by Mano et al. [76], who conducted a study comparing various ML models for predicting virological failure in HIV patients. In their study, RF demonstrated the highest accuracy [76]. Deep learning models such as RNN, PLS-DA, and CNN (MobileNet2) were also identified as the best-performing algorithms in a few studies [27,40–42]. Likewise, SVM, LR, GCN, and monogram outperformed other models in some studies. Several factors influence the predictive performance of ML models, including data size, missing values, feature and engineer selection, ensemble methods, cross-validation, and the expertise of the person conducting the analysis [77,78]. Thus, it might be difficult to point out which algorithms are more accurate than others, despite the complexity of certain models.

4.2. Successes and Opportunities of Machine Learning Interventions in HIV Testing

This study demonstrated that algorithms can effectively predict individuals at the highest risk of HIV as well as crucial related parameters to determine the priority population for HIV testing. Our findings indicate that ML methods outperformed human efforts and traditional statistical predictive models utilized across countries to improve HIV testing. The independent *t*-test from our study shows that ML models achieved 12.53% ($p = 0.002$) higher accuracy compared to TMs. These findings are consistent with those of Fieggen et al. [6], who reviewed different ML models and concluded that they were more effective in HIV risk prediction than traditional approaches.

Several studies forecast the number of people who will be HIV-positive in the following decade [15]. In addition, socio-behavioral data have been used to predict HIV incidence and prevalence in the Philippines and a few African countries [51,53]. This is critical for the HIV testing cascade because it enables relevant stakeholders to identify which groups of people are most at risk of contracting HIV and to plan targeted interventions for efficient resource allocation [10]. Findings from a review of AI applications in HIV prevention by Marcus et al. [69] align with ours. Moreover, the study uncovers the power of ML in predicting which socio-demographic and socio-cultural factors are associated with HIV testing. In the USA, old age, inpatient visits, alcohol use, male gender, condom use, sexual behavior, and substance use were identified as key predictors of HIV risk and diagnosis [19–22]. Similar to the finding of a study that looked at gender differences as predictors of sexual behaviors and HIV testing in Kenya, young men were found to be more likely to engage in risky sexual behavior [79]. Despite this, males are more reluctant to undertake HIV testing due to barriers such as knowledge deficits, stigma, and perceived HIV risk [80,81]. However, Laybohr Kamara et al. [53] reveal that HIV prevalence is higher in women in the Mano River countries. Our results are consistent with those of a review of ML in HIV prevention by the authors of [69], whose findings indicate that ML has the potential to identify individuals at the highest risk of HIV for the early initiation of treatment.

In addition, KPs such as MSM are one of society's most vulnerable groups. Even in areas with a supportive legal framework, they face societal stigma. This impairs their willingness to seek medical attention, exposing them and their partners to HIV risk. This study demonstrates the predictive power of ML interventions for HIV risk, testing, and status. Asymptomatic screening of STI symptoms, sex overseas in the previous 12 months, persons living with HIV, contact with infection, current sex work, screening reminder frequency, the type of screening reminder method, sex with a male in the previous three months, and condom use were among the major predictors of HIV risk, testing, and status among MSM, according to studies from Australia and China [24,32].

This review affirms that ML interventions can enhance accurate and efficient HIV testing. They are not only capable of predicting the prevalence of HIV testing, but are also applicable in diagnostic devices to detect false-positive HIV test results and improve the privacy, accessibility, and convenience of HIV testing. With the aid of DL in France, 38% of false-positive HIV results were detected from a retrospective analysis [25]. Likewise, findings from the USA detected a significant number of false-positive HIV test results using

ML models [22,57]. This represents a significant milestone in pursuing the 95-95-95 goal, as the use of ML prevents testers from experiencing conflicting emotions and strengthens their confidence in healthcare providers. Machine learning has also been very effective in reducing false-negative HIV test results, which are more concerning because falsely diagnosed PLHIV can increase the spread of HIV [48,55]. It is also quite thrilling to discover an AI conversational agent built in South Africa to enhance self-automated HIV testing with evidence of improvement in anonymity, privacy, speed, and accessibility [17]. Other AI-based mobile algorithms (mHealth and HealthPulse) effectively detect and interpret RDTs for HIV, while prioritizing privacy, accountability, and accessibility in LMICs [48,50,55].

4.3. Gaps and Challenges Identified in the Implementation of Machine Learning in HIV Testing

Although our analysis suggests that ML is generally more accurate than traditional HIV testing methods, it is important to note that ML algorithms are not always infallible and may not constantly outperform traditional techniques. In addition, there exist limitations, challenges, and gaps in implementing ML interventions, especially in resource-constrained countries. The study's findings reported that biases were encountered in certain training data, which caused unrepresentative results in some studies. Furthermore, most studies examined in this review utilized retrospective designs, incorporating self-reported information vulnerable to missing data and recall bias. These issues can also result in skewed models, potentially leading to inaccurate predictions and disparities in HIV testing outcomes. Our finding is consistent with a statistical review by Zhou et al. [82], who acknowledged that ML techniques have several advantages over traditional statistical methods, but that the use of ML algorithms poses numerous challenges. In their study, bias in unstructured data, bias algorithms, and model overfitting were deemed limitations that could impair the effective predictive performance of ML models, resulting in erroneous outcomes [82]. Reinforcing this finding, our analysis also came across underfitted and overfitted models and several studies that suffered from small sample sizes and lacked representativeness [17,20,29,48,60].

Diversity in geographical locations also affects the generalization of models. This study highlighted that models employed in a specific region or among a particular demographic may not be transferable to other regions or populations. Moreover, insufficient expertise and the usability issues of certain models are hampering the universal applicability of ML in HIV testing interventions, especially in underprivileged settings. This explains why ML approaches are more common in developed countries. We found that most ML initiatives were concentrated in South Africa because of the region's recent technological advancements catalyzed by the intense fight against HIV.

To achieve the UNAIDS 95-95-95 goal of eradicating HIV/AIDS by 2030 with ML interventions, certain areas need to be improved. Hence, more ML studies should be applied in prospective studies to yield more meaningful data that can accurately predict individual behavior risk without recall bias. There is also a need for more professional training in ML for improved usability and acceptability, especially in resource-constrained settings. Furthermore, more studies involving ML in HIV testing interventions should be conducted across SSA, especially in high-HIV-incidence epidemic settings [32].

4.4. Strengths and Limitations of the Study

This study followed a comprehensive review process, gaining deep insights into the extent of ML usage, its associated benefits, and its limitations. At first, the protocol was prospectively registered with PROSPERO to avoid duplication. In addition, the PRISMA guidelines were followed to generate consistent results and enhance the study's replicability. A comprehensive literature search was conducted in a global context across different journal databases. The Covidence software aided in the exclusion of duplicates and ineligible articles, and was quite fast and reliable. Two reviewers screened the articles, and a third party was involved in resolving conflicts regarding the articles to include in the analysis. We conducted a *t*-test to determine the accuracy of ML over TMs.

However, a few challenges were encountered during the review process. This review excluded studies that used ML algorithms to predict HIV testing among individuals less than 15 years old, which limits the generalizability of our findings to only those that are 15 years or older. Additionally, the researchers could not access certain subscription-based articles, which led to the exclusion of some relevant articles. Furthermore, the TMs and ML models differ, as shown by the accuracies in eight studies that applied TMs, which were compared to those of 36 ML models. A normality test was conducting before running an independent *t*-test to ensure that the data were normally distributed. The *t*-test results show a statistically significant difference ($p = 0.0002$) between ML and TMs. Future studies should investigate the application of ML HIV testing interventions in age categories below 15 years.

5. Conclusions

Early detection and diagnosis of HIV is a starting point for achieving the UNAIDS 95-95-95 goal of ending AIDS as a global pandemic by 2030. However, rapid population growth coupled with emerging diseases and resource constraints, especially in developing settings, makes this goal unattainable with conventional HIV testing interventions. It is critical to provide new HIV testing strategies for specific individuals/high-risk populations to maximize resource allocation in resource-constrained countries. This study provides valuable insights into the current state of research at the intersection of ML and HIV testing. It points to the positive impact of ML in enhancing the early prediction of HIV spread, optimizing HIV testing approaches, improving efficiency, and ultimately enhancing the accuracy of HIV diagnosis. This study further reveals that ML techniques are more accurate than traditional approaches in predicting HIV risk, testing, and status. Also, most ML interventions are concentrated in developed countries due to a lack of expertise and the inapplicability of models. Research institutions should train more HIV epidemiologists to become ML experts. Screening programs should incorporate automated HIV testing models for improved privacy, acceptability, and accessibility with high diagnostic accuracy.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/a17080362/s1: Table S1: PICOS framework; Table S2: Data extraction tool; Table S3: CASPSR checklist; Table S4: PRISMA checklist [83]; Table S5: Study characteristics; Table S6: Machine learning interventions applied; Table S7: Strengths and limitations.

Author Contributions: M.J. wrote the first draft of the paper. M.J., R.N.P.-M., E.P., Y.A.S. and K.A.V. conceived the study and provided insightful views on its conception. R.N.P.-M., E.P., Y.A.S. and K.A.V. reviewed the work. R.N.P.-M., E.P. and Y.A.S. supervised the entire study. M.J., R.N.P.-M. and E.P. reviewed the literature, screened all articles, and conducted quality appraisal. M.J. wrote the introduction, methods, data abstraction, results, narrative analysis, and discussion sections. All authors have read and agreed to the published version of the manuscript.

Funding: The work reported herein was made possible through funding by the South African Medical Research Council (SAMRC) through its Division of Research Capacity Development under the Mid-Career Scientist Program using funding received from the South African National Treasury (Project Code number: 57035 [SAMRC File ref no: HDID8528/KR/202]). The authors Refilwe Nancy Phaswana-Mafuya, Edith Phalane, and Musa Jaiteh are supported by the above-mentioned grant. This work was conducted under the guidance of SAMRC/University of Johannesburg (UJ) Pan African Centre for Epidemics Research (PACER) Extramural Unit. The content herein is the authors' sole responsibility and does not necessarily represent the official views of SAMRC or UJ. This paper also forms part of a PhD study by Mr Musa Jaiteh, whose studies are funded by the Global Excellence Stature 4.0 Scholarship at UJ.

Data Availability Statement: Not applicable.

Acknowledgments: The study presented in this publication falls under the Boloka Project, which is funded by the SAMRC under its Division of Research Capacity Development through the Mid-Career Scientist Program, utilizing funds allocated by the South African National Treasury. The authors Refilwe Nancy Phaswana-Mafuya, Edith Phalane, and Musa Jaiteh are supported by SAMRC. This systematic review also forms part of a doctoral study by Musa Jaiteh, whose studies are funded by the UJ Global Excellence Stature (GES) 4.0 Scholarship. It is conducted within the SAMRC/UJ PACER Extramural Unit framework. The views expressed in this work are the authors' sole responsibility and do not necessarily reflect the official perspectives of the SAMRC or UJ.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. World Health Organization. H.I.V. 2024. Available online: https://www.who.int/data/gho/data/themes/hiv-aids (accessed on 11 February 2024).
2. World Health Organization. Consolidated Guidelines on HIV Prevention, Testing, Treatment, Service Delivery, and Monitoring: Recommendations for a Public Health Approach. 2021. Available online: https://www.who.int/publications/i/item/9789240031593 (accessed on 17 February 2024).
3. UNAIDS. *Prevailing against Pandemics by Putting People at the Centre—World AIDS Day Report 2020*; UNAIDS: Geneva, Switzerland, 2020; pp. 1–92. Available online: https://www.unaids.org/sites/default/files/media_asset/prevailing-against-pandemics_en.pdf (accessed on 17 February 2024).
4. UNAIDS. *FACT SHEET 2023 Global HIV Statistics People Living with HIV People Living with HIV Accessing Antiretroviral Therapy New HIV Infections AIDS-Related Deaths Key Populations*; UNAIDS: Geneva, Switzerland, 2023.
5. Mannoh, I.; Amundsen, D.; Turpin, G.; Lyons, C.E.; Viswasam, N.; Hahn, E.; Ryan, S.; Baral, S.; Hansoti, B. A Systematic Review of HIV Testing Implementation Strategies in Sub-Saharan African Countries. *AIDS Behav.* **2022**, *26*, 1660–1671. [CrossRef] [PubMed]
6. Fieggen, J.; Smith, E.; Arora, L.; Segal, B. The role of machine learning in HIV risk prediction. *Front. Reprod. Health* **2022**, *4*, 1062387. [CrossRef]
7. World Health Organization. WHO Recommends HIV Self-Testing—Evidence Update and Considerations for Success. 2024. Available online: https://www.who.int/publications/i/item/WHO-CDS-HIV-19.36#:~:text=In%202016%20WHO%20recommended%20HIVST%20as%20a%20safe,,performance%20comparable%20to%20that%20of%20trained%20health-care%20workers (accessed on 14 July 2024).
8. Van Staden, Q.; Laurenzi, C.A.; Toska, E. Two years after lockdown: Reviewing the effects of COVID-19 on health services and support for adolescents living with HIV in South Africa. *J. Int. AIDS Soc.* **2022**, *25*, e25904. [CrossRef]
9. 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*, e54–e62. [CrossRef] [PubMed]
10. Mutai, C.K.; McSharry, P.E.; Ngaruye, I.; Musabanganji, E. Use of machine learning techniques to identify HIV predictors for screening in sub-Saharan Africa. *BMC Med. Res. Methodol.* **2021**, *21*, 159. [CrossRef] [PubMed]
11. Dubey, A. Biophotonics and machine learning model for the diagnosis and treatment of HIV. *Biosci. Biotechnol. Res. Commun.* **2018**, *11*, 5–10. [CrossRef]
12. Turk, G.; Ghiglione, Y.; Hormanstorfer, M.; Laufer, N.; Coloccini, R.; Salido, J.; Trifone, C.; Ruiz, M.J.; Falivene, J.; Holgado, M.P.; et al. Biomarkers of progression after HIV acute/early infection: Nothing compares to CD4+ T-cell count? *Viruses* **2018**, *10*, 34. [CrossRef]
13. Chikusi, H.; Leo, J.; Kaijage, S. Machine Learning Model for Prediction and Visualization of HIV Index Testing in Northern Tanzania. *Int. J. Adv. Comput. Sci. Appl.* **2022**, *13*, 391–399. [CrossRef]
14. Oladokun, O.O. *Predicting HIV Status among Women in South Africa Using Machine Learning: Comparing Decision Tree Model and Logistic Regression*; WIReDSpace: Johannesburg, South Africa, 2019; Available online: https://wiredspace.wits.ac.za/bitstreams/d6c6cc37-2e79-4f6a-a35c-b978ed067842/download (accessed on 14 July 2024).
15. Ahlström, M.G.; Ronit, A.; Omland, L.H.; Vedel, S.; Obel, N. Algorithmic prediction of HIV status using nation-wide electronic registry data. *EClinicalMedicine* **2019**, *17*, 100203. [CrossRef]
16. Alehegn, M. Application of machine learning and deep learning for the prediction of HIV/AIDS. *HIV AIDS Rev.* **2022**, *21*, 17–23. [CrossRef]
17. Van Heerden, A.; Ntinga, X.; Vilakazi, K. The potential of conversational agents to provide a rapid HIV counseling and testing services. In Proceedings of the 2017 International Conference on the Frontiers and Advances in Data Science (FADS), Xi'an, China, 23–25 October 2017; IEEE: Piscataway, NJ, USA, 2018; pp. 80–85.
18. Wood, C.S.; Thomas, M.R.; Budd, J.; Mashamba-Thompson, T.P.; Herbst, K.; Pillay, D.; Peeling, R.W.; Johnson, A.M.; McKendry, R.A.; Stevens, M.M. Taking connected mobile-health diagnostics of infectious diseases to the field. *Nature* **2019**, *566*, 467–474. [CrossRef] [PubMed]
19. Pan, Y.; Liu, H.; Metsch, L.R.; Feaster, D.J. Factors Associated with HIV Testing Among Participants from Substance Use Disorder Treatment Programs in the US: A Machine Learning Approach. *AIDS Behav.* **2017**, *21*, 534–546. [CrossRef]
20. Ovalle, A.; Goldstein, O.; Kachuee, M.; Wu, E.S.C.; Hong, C.; Holloway, I.W.; Sarrafzadeh, M. Leveraging social media activity and machine learning for HIV and substance abuse risk assessment: Development and validation study. *J. Med. Internet Res.* **2021**, *23*, e22042. [CrossRef]

21. Xiang, Y.; Fujimoto, K.; Schneider, J.; Jia, Y.; Zhi, D.; Tao, C. Network context matters: Graph convolutional network model over social networks improves the detection of unknown HIV infections among young men who have sex with men. *J. Am. Med. Informatics Assoc.* **2019**, *26*, 1263–1271. [CrossRef]
22. Elkhadrawi, M.; Bryan, A.S.; Bradley, J.W.; Murat, A.S.W. Challenges in communication from referring clinicians to pathologists in the electronic health record era. *J. Pathol. Inform.* **2021**, *9*, 8.
23. Burns, C.M.; Pung, L.; Witt, D.; Gao, M.; Sendak, M.; Balu, S.; Krakower, D.; Marcus, J.L.; Okeke, N.L.; E Clement, M. Development of a Human Immunodeficiency Virus Risk Prediction Model Using Electronic Health Record Data From an Academic Health System in the Southern United States. *Clin. Infect. Dis.* **2023**, *76*, 299–306. [CrossRef]
24. Xu, X.; Fairley, C.K.; Chow, E.P.F.; Lee, D.; Aung, E.T.; Zhang, L.; Ong, J.J. Using machine learning approaches to predict timely clinic attendance and the uptake of HIV/STI testing post clinic reminder messages. *Sci. Rep.* **2022**, *12*, 8757. [CrossRef]
25. Demey, B.; Usureau, C.; Boillod, P.; Bodeau, S.; François, C.; Roussel, C.; Duverlie, G.; Castelain, S.; Brochot, E. Analytical performance evaluation and enhancement of the ADVIA Centaur® HIV Ag/Ab Combo assay. *J. Clin. Virol.* **2019**, *118*, 36–40. [CrossRef]
26. Equator Network. The PRISMA 2020 Statement: The Updated Guideline for Reporting Systematic Reviews. 2024. Available online: https://www.equator-network.org/reporting-guidelines/prisma/ (accessed on 18 April 2024).
27. Amir-Behghadami, M.; Janati, A. Population, Intervention, Comparison, Outcomes and Study (PICOS) design as a framework to formulate eligibility criteria in systematic reviews. *Emerg. Med. J.* **2020**, *37*, 387. [CrossRef] [PubMed]
28. RefWorks. RefWorks. Making Research Easier. 2024. Available online: https://refworks.proquest.com/ (accessed on 22 February 2024).
29. Saha, R.; Malviya, L.; Jadhav, A.; Dangi, R. Early stage HIV diagnosis using optimized ensemble learning technique. *Biomed. Signal Process. Control.* **2024**, *89*, 105787. [CrossRef]
30. Cobre, A.; Morais, A.; Selege, F.; Stremel, D.; Wiens, A.; Ferreira, L.; Tonin, F.; Pontarolo, R. Use of Biochemical Tests and Machine Learning in the Search for Potential Diagnostic Biomarkers of COVID-19, HIV/AIDS, and Pulmonary Tuberculosis. *J. Braz. Chem. Soc.* **2024**, *35*, e-20240020. [CrossRef]
31. Alie, M.S.; Negesse, Y. Machine learning prediction of adolescent HIV testing services in Ethiopia. *Front. Public Health* **2024**, *12*, 1341279. [CrossRef]
32. He, J.; Li, J.; Jiang, S.; Cheng, W.; Jiang, J.; Xu, Y.; Yang, J.; Zhou, X.; Chai, C.; Wu, C. Application of machine learning algorithms in predicting HIV infection among men who have sex with men: Model development and validation. *Front. Public Health* **2022**, *10*, 967681. [CrossRef]
33. Balzer, L.B.; Havlir, D.V.; Kamya, M.R.; Chamie, G.; Charlebois, E.D.; Clark, T.D.; A Koss, C.; Kwarisiima, D.; Ayieko, J.; Sang, N.; et al. Machine Learning to Identify Persons at High-Risk of Human Immunodeficiency Virus Acquisition in Rural Kenya and Uganda. *Clin. Infect. Dis.* **2020**, *71*, 2326–2333. [CrossRef]
34. Bao, Y.; Medland, N.A.; Fairley, C.K.; Wu, J.; Shang, X.; Chow, E.P.F.; Xu, X.; Ge, Z.; Zhuang, X.; Zhang, L. Predicting the diagnosis of HIV and sexually transmitted infections among men who have sex with men using machine learning approaches. *J. Infect.* **2021**, *82*, 48–59. [CrossRef]
35. Birri Makota, R.B.; Musenge, E. Predicting HIV infection in the decade (2005–2015) pre-COVID-19 in Zimbabwe: A supervised classification-based machine learning approach. *PLOS Digit. Health* **2023**, *2*, e0000260. [CrossRef]
36. Chingombe, I.; Musuka, G.; Mbunge, E.; Chemhaka, G.; Cuadros, D.F.; Murewanhema, G.; Chaputsira, S.; Batani, J.; Muchemwa, B.; Mapingure, M.P.; et al. Predicting HIV Status Using Machine Learning Techniques and Bio-Behavioural Data from the Zimbabwe Population-Based HIV Impact Assessment (ZIMPHIA15-16). In *LNNS, Lecture Notes in Networks and Systems*; Springer International Publishing: Cham, Switzerland, 2022; Volume 502, pp. 247–258. [CrossRef]
37. Dong, Y.; Liu, S.; Xia, D.; Xu, C.; Yu, X.; Chen, H.; Wang, R.; Liu, Y.; Dong, J.; Hu, F.; et al. Prediction Model for the Risk of HIV Infection among MSM in China: Validation and Stability. *Int. J. Environ. Res. Public Health* **2022**, *19*, 1010. [CrossRef]
38. Friedman, E.E.; Shankaran, S.; Devlin, S.A.; Kishen, E.B.; Mason, J.A.; Sha, B.E.; Ridgway, J.P. Development of a predictive model for identifying women vulnerable to HIV in Chicago. *BMC Womens Health* **2023**, *23*, 313. [CrossRef] [PubMed]
39. Jing, F.; Ye, Y.; Zhou, Y.; Ni, Y.; Yan, X.; Lu, Y. Identification of Key Influencers for Secondary Distribution of HIV Self-Testing Kits Among Chinese Men Who Have Sex With Men: Development of an Ensemble Machine Learning Approach. *J. Med. Internet Res.* **2023**, *25*, e37719. [CrossRef] [PubMed]
40. Orel, E.; Esra, R.; Estill, J.; Thiabaud, A.; Marchand-Maillet, S.; Merzouki, A.; Keiser, O. Prediction of HIV status based on sociobehavioural characteristics in East and Southern Africa. *PLoS ONE* **2022**, *17*, e0264429. [CrossRef]
41. Ni, Y.; Lu, Y.; Jing, F.; Wang, Q.; Xie, Y.; He, X. A Machine Learning Model for Identifying Sexual Health Influencers to Promote the Secondary Distribution of HIV Self-Testing Among Gay, Bisexual, and Other Men Who Have Sex With Men in China: Quasi-Experimental Study. *JMIR Public Health Surveill.* **2024**, *10*, e50656. [CrossRef]
42. Majam, M.; Segal, B.; Fieggen, J.; Smith, E.; Hermans, L.; Singh, L.; Phatsoane, M.; Arora, L.; Lalla-Edward, S. Informatics in Medicine Unlocked Utility of a machine-guided tool for assessing risk behaviour associated with contracting HIV in three sites in South Africa. *Inform. Med. Unlocked* **2023**, *37*, 101192. [CrossRef]
43. Hoenigl, M.; Weibel, N.; Mehta, S.R.; Anderson, C.M.; Jenks, J.; Green, N.; Gianella, S.; Smith, D.M.; Little, S.J. Development and Validation of the San Diego Early Test Score to Predict Acute and Early HIV Infection Risk in Men Who Have Sex With Men. *Clin. Infect. Dis.* **2015**, *61*, 468–475. [CrossRef] [PubMed]

44. Hoots, B.E.; MacDonald, P.D.; Hightow-Weidman, L.B.; Leone, P.A.; Miller, W.C. Developing a predictive model to prioritize human immunodeficiency virus partner notification in North Carolina. *Sex. Transm. Dis.* **2012**, *39*, 65–71. [CrossRef]
45. Wade, B.S.C.; Valcour, V.G.; Wendelken-riegelhaupt, L.; Esmaeili-firidouni, P.; Joshi, S.H.; Gutman, B.A.; Thomson, P.M. Clinical Mapping abnormal subcortical brain morphometry in an elderly HIV + cohort. *YNICL* **2015**, *9*, 564–573. [CrossRef]
46. Van Nguyen, T.P.; Yang, W.; Tang, Z.; Xia, X.; Mullens, A.B.; Dean, J.A.; Li, Y. Lightweight federated learning for STIs/HIV prediction. *Sci. Rep.* **2024**, *14*, 6560. [CrossRef]
47. Chingombe, I.; Dzinamarira, T.; Cuadros, D.; Mapingure, M.P.; Mbunge, E.; Chaputsira, S.; Madziva, R.; Chiurunge, P.; Samba, C.; Herrera, H.; et al. Predicting HIV Status among Men Who Have Sex with Men in Bulawayo & Harare, Zimbabwe Using Bio-Behavioural Data, Recurrent Neural Networks, and Machine Learning Techniques. *Trop. Med. Infect. Dis.* **2022**, *7*, 231. [CrossRef]
48. Turbé, V.; Herbst, C.; Mngomezulu, T.; Meshkinfamfard, S.; Dlamini, N.; Mhlongo, T.; Smit, T.; Cherepanova, V.; Shimada, K.; Budd, J.; et al. Deep learning of HIV field-based rapid tests. *Nat. Med.* **2021**, *27*, 1165–1170. [CrossRef]
49. Belete, D.; Huchaiah, M.D. A Deep Learning Approaches for Modeling and Predicting of HIV Test Results Using EDHS Dataset. In *Infectious Diseases*; IntechOpen: London, UK, 2023. [CrossRef]
50. Gupta, K.; Ruan, Y.; Ibrahim, A.; Mendonca, R.; Cooper, S.; Morris, S.; Hattery, D. Transforming Rapid Diagnostic Tests into Trusted Diagnostic Tools in LMIC using AI. In Proceedings of the 2023 IEEE Conference on Artificial Intelligence (CAI), Santa Clara, CA, USA, 5–6 June 2023; IEEE: Piscataway, NJ, USA, 2023; pp. 306–308.
51. Aribe, S.G.; Gerardo, B.D.; Medina, R.P. Time Series Forecasting of HIV/AIDS in the Philippines Using Deep Learning: Does COVID-19 Epidemic Matter? *Int. J. Emerg. Technol. Adv. Eng.* **2022**, *12*, 144–157.
52. Bender Ignacio, R.A.; Dasgupta, S.; Valdez, R.; Pandey, U.; Pasalar, S.; Alfaro, R.; Hladik, F.; Gornalusse, G.; Lama, J.R.; Duerr, A. Dynamic immune markers predict HIV acquisition and augment associations with sociobehavioral factors for HIV exposure. *iScience* **2022**, *25*, 105632. [CrossRef]
53. Laybohr Kamara, I.; Wang, L.; Guo, Y.; Huo, S.; Guo, Y.; Xu, C.; Liao, Y.; Liu, W.J.; Ma, W.; Gao, G.F. Spatial–temporal heterogeneity and determinants of HIV prevalence in the Mano River Union countries. *Infect. Dis. Poverty.* **2022**, *11*, 116. [CrossRef]
54. Wang, X.; Yang, M.; Li, W. Efficient Data Reduction Strategies for Big Data and High-Dimensional LASSO Regressions. *arXiv* **2024**, arXiv:2401.11070. [CrossRef]
55. Roche, S.D.; Ekwunife, O.I.; Mendonca, R.; Kwach, B.; Omollo, V.; Zhang, S.; Ongwen, P.; Hattery, D.; Smedinghoff, S.; Morris, S.; et al. Measuring the performance of computer vision artificial intelligence to interpret images of HIV self-testing results. *Front. Public Health* **2024**, *12*, 1334881. [CrossRef]
56. Latt, P.M.; Soe, N.N.; Xu, X.; Ong, J.J.; Chow, E.P.F.; Fairley, C.K.; Zhang, L. Identifying Individuals at High Risk for HIV and Sexually Transmitted Infections With an Artificial Intelligence-Based Risk Assessment Tool. *Open Forum Infect. Dis.* **2024**, *11*, ofae011. [CrossRef] [PubMed]
57. Zucker, J.; Carnevale, C.; Gordon, P.; Sobieszczyk, M.E.; Rai, A.J. Am i Positive? Improving Human Immunodeficiency Virus Testing in the Era of Preexposure Prophylaxis and Immediate Antiretroviral Therapy Using Machine Learning. *Open Forum Infect. Dis.* **2022**, *9*, ofac259. [CrossRef]
58. Mutai, C.K.; McSharry, P.E.; Ngaruye, I.; Musabanganji, E. Use of unsupervised machine learning to characterise HIV predictors in sub-Saharan Africa. *BMC Infect. Dis.* **2023**, *23*, 482. [CrossRef]
59. Feller, D.J.; Zucker, J.; Yin, M.T.; Gordon, P.; Elhadad, N. Using Clinical Notes and Natural Language Processing for Automated HIV Risk Assessment. *J. Acquir. Immune Defic. Syndr.* **2018**, *77*, 160–166. [CrossRef]
60. Cheah, M.H.; Gan, Y.N.; Altice, F.L.; Wickersham, J.A. Testing the Feasibility and Acceptability of Using an Artificial Intelligence Chatbot to Promote HIV Testing and Pre-Exposure Prophylaxis in Malaysia: Mixed Methods Study. *JMIR Hum. Factors* **2024**, *11*, e52055. [CrossRef]
61. Xu, X.; Yu, Z.; Ge, Z.; Chow, E.P.F.; Bao, Y.; Ong, J.J.; Li, W.; Wu, J.; Fairley, C.K.; Zhang, L. Web-Based Risk Prediction Tool for an Individual's Risk of HIV and Sexually Transmitted Infections Using Machine Learning Algorithms: Development and External Validation Study. *J. Med. Internet Res.* **2022**, *24*, e37850. [CrossRef]
62. Haas, O.; Maier, A.; Rothgang, E. Machine Learning-Based HIV Risk Estimation Using Incidence Rate Ratios. *Front. Reprod. Health* **2021**, *3*, 1–10. [CrossRef]
63. Jing, F.; Zhang, Q.; Ong, J.J.; Xie, Y.; Ni, Y.; Cheng, M.; Huang, S.; Zhou, Y.; Tang, W. Optimal resource allocation in HIV self-testing secondary distribution among Chinese MSM: Data-driven integer programming models. *Philos. Trans. R. Soc. A* **2021**, *380*, 20210128. [CrossRef] [PubMed]
64. Scott, H.; Vittinghoff, E.; Irvin, R.; Liu, A.; Nelson, L.; Del, C.; Magnus, M.; Mannheimer, S.; Fields, S.; Van Tieu, H.; et al. Development and Validation of the Personalized Sexual Health Promotion (SexPro) HIV Risk Prediction Model for Men Who Have Sex with Men in the United States. *AIDS Behav.* **2020**, *24*, 274–283. [CrossRef]
65. Adeli, E.; Zahr, N.M.; Pfefferbaum, A.; Sullivan, E.V.; Pohl, K.M. Novel Machine Learning Identifies Brain Patterns Distinguishing Diagnostic Membership of Human Immunodeficiency Virus, Alcoholism, and Their Comorbidity of Individuals. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* **2019**, *4*, 589–599. [CrossRef]
66. Rice, E.; Yoshioka-Maxwell, A.; Petering, R.; Onasch-Vera, L.; Craddock, J.; Tambe, M.; Yadav, A.; Wilder, B.; Woo, D.; Winetrobe, H.; et al. Piloting the Use of Artificial Intelligence to Enhance HIV Prevention Interventions for Youth Experiencing Homelessness. *J. Soc. Soc. Work. Res.* **2018**, *9*, 551–573. [CrossRef]

67. Menza, T.W.; Hughes, J.P.; Celum, C.L.; Golden, M.R. Prediction of HIV acquisition among men who have sex with men. *Sex. Transm. Dis.* **2009**, *36*, 547–555. [CrossRef]
68. Weissman, S.; Yang, X.; Zhang, J.; Chen, S.; Olatosi, B.; Li, X. Using a machine learning approach to explore predictors of healthcare visits as missed opportunities for HIV diagnosis. *AIDS* **2021**, *35* (Suppl. S1), S7–S18. [CrossRef]
69. Marcus, J.L.; Sewell, W.C.; Balzer, L.B.; Krakower, D.S. Artificial Intelligence and Machine Learning for HIV Prevention: Emerging Approaches to Ending the Epidemic. *Curr. HIV/AIDS Rep.* **2020**, *17*, 171–179. [CrossRef]
70. Rasheed, K.; Qayyum, A.; Ghaly, M.; Al-Fuqaha, A.; Razi, A.; Qadir, J. Explainable, trustworthy, and ethical machine learning for healthcare: A survey. *Comput. Biol. Med.* **2022**, *149*, 106043. [CrossRef]
71. Lainjo, B. Artificial intelligence with machine learning and the enigmatic discovery of HIV cure. *J. Auton. Intell.* **2024**, *7*, 1–16. [CrossRef]
72. Runde, D.F.; Bandura, R.; Ramanujam, S. *The United States Has an Opportunity to Lead in Digital Development*; Center for Strategic & International Studies: Washington, DC, USA, 2021; pp. 1–17. Available online: https://www.csis.org/analysis/united-states-has-opportunity-lead-digital-development (accessed on 5 July 2024).
73. Pythonista Planet. Pros and Cons of Supervised Machine Learning. 2024. Available online: https://pythonistaplanet.com/pros-and-cons-of-supervised-machine-learning/#google_vignette (accessed on 1 February 2024).
74. Rogers. Captera. Supervised vs Unsupervised Learning: Which Machine Learning Model Is Right for You? 2021. Available online: https://www.capterra.com/resources/supervised-vs-unsupervised-learning/ (accessed on 1 February 2024).
75. Ferreira, L.; Pilastri, A.; Martins, C.M.; Pires, P.M.; Cortez, P. A Comparison of AutoML Tools for Machine Learning, Deep Learning and XGBoost. In Proceedings of the 2021 International Joint Conference on Neural Networks (IJCNN), Shenzhen, China, 1–8 July 2021.
76. Mamo, D.N.; Yilma, T.M.; Fekadie, M.; Sebastian, Y.; Bizuayehu, T.; Melaku, M.S.; Walle, A.D. 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.* **2023**, *23*, 75. [CrossRef]
77. Ray, S. *Ways to Improve Accuracy of Machine Learning Models (Updated 2024)*; Analytics Vidhya: Gurugram, India, 2024.
78. Bognár, L.; Fauszt, T. Factors and conditions that affect the goodness of machine learning models for predicting the success of learning. *Comput. Educ. Artif. Intell.* **2022**, *3*, 100100. [CrossRef]
79. Chenneville, T.; Drake, H. Gender Differences in Psychosocial Predictors of Sexual Activity and HIV Testing Among Youth in Kenya. *Front. Reprod. Health* **2021**, *3*, 636462. [CrossRef]
80. Kalichman, S.C.; Shkembi, B.; Wanyenze, R.K.; Naigino, R.; Bateganya, M.H.; Menzies, N.A.; Lin, C.-D.; Lule, H.; Kiene, S.M. Perceived HIV stigma and HIV testing among men and women in rural Uganda: A population-based study. *Lancet HIV* **2020**, *7*, e817–e824. [CrossRef]
81. Hlongwa, M.; Mashamba-Thompson, T.; Makhunga, S.; Hlongwana, K. Barriers to HIV testing uptake among men in sub-Saharan Africa: A scoping review. *Afr. J. AIDS Res.* **2020**, *19*, 13–23. [CrossRef]
82. Zhou, N.; Zhang, Z.; Nair, V.N.; Singhal, H.; Chen, J. Bias, Fairness and Accountability with Artificial Intelligence and Machine Learning Algorithms. *Int. Stat. Rev.* **2022**, *90*, 468–480. [CrossRef]
83. Page, M.J.; McKenzie, J.E.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E.; et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ* **2021**, *372*, n71.