GSC Biological and Pharmaceutical Sciences

eISSN: 2581-3250 CODEN (USA): GBPSC2

Cross Ref DOI: 10.30574/gscbps

Journal homepage: https://gsconlinepress.com/journals/gscbps/

(REVIEW ARTICLE)

Check for updates

Machine learning and pre-exposure prophylaxis: A survey

Judith N. Nyakanga *

Uzima University, Kisumu, Kenya.

GSC Biological and Pharmaceutical Sciences, 2023, 25(03), 005–014

Publication history: Received on 03 May 2023; revised on 30 November 2023; accepted on 03 December 2023

Article DOI: https://doi.org/10.30574/gscbps.2023.25.3.0182

Abstract

The goal of this paper was to study how machine learning techniques have been applied in PrEP (Pre-exposure prophylaxis) and HIV prediction to identify individuals who are at a higher risk of acquiring HIV infection and to optimize the used of PrEP, which is an effective method of preventing HIV transmission. The results indicate that machine learning has been used to HIV risk, optimizing PrEP use, developing personalized PrEP regimens, and identifying PrEP candidates. In predicting HIV risk, machine learning algorithms have been developed to predict HIV risk based on various factors such as demographic information, sexual behavior and drug use. The models can identify individuals who are at a higher risk of acquiring HIV infection and can be used to target interventions such as PrEP to those who are most in need. Regarding optimizing PrEP use, machine learning has been utilized to optimize PrEP usage by identifying the factors that are associated with adherence to PrEP. These models can help healthcare providers to tailor their interventions to promote PrEP adherence and improve its effectiveness. In addition, machine learning techniques have been used to develop personalized PrEP regimens based on an individual's HIV risk profile. These models can help healthcare providers to optimize PrEP use and reduce the risk of HIV transmission. It was also established that machine learning models have been used to identify individuals who are most likely to benefit from PrEP. These models can help healthcare providers to target PrEP interventions to those who are likely to benefit from them.

Keywords: AIDS; Machine learning; HIV; Prediction; PrEP

1. Introduction

The HIV epidemic is evaluated to affect 76 million people worldwide, roughly 33 million of whom have died [1]. Whereas there is a 60% reduction in evaluated AIDS-related annual deaths, the development has not been considered in HIV incidence which have a 17% decrease in HIV incidence. This shows an increase in the number of people living with HIV [2], [3]. UNAIDS target of 95-95-95 is the fast track for reducing HIV incidences to achieve the global efforts to end HIV Pandemic by 2030. This means that 95% of people living with HIV should know their status, 95 % of the diagnosed put on the antiretroviral and 95% of those on antiretrovirals have viral suppression by 2030. Although there are advances in reducing HIV mortality, there are high incidence rates that are averting the global attempt to end the pandemic by 2030. Table 1 provides examples of major breakthroughs in the field of PrEP and HIV prediction.

UNAIDS has laid out some preventive strategies like PrEP (Pre-exposure prophylaxis) as one of the priorities to lessen the transmission that is undermining the global efforts to end the HIV pandemic. The authors in [4] suggest that accurate and granular risk prediction is crucial for the campaigns, though they may be insufficient in areas where the burden is soaring. PrEP and antiretroviral drugs have been shown to be 100% effective when taken correctly.

* Corresponding author: Judith N. Nyakanga

Table 1 Major breakthroughs in PrEP and HIV prediction

Breakthrough	Description
Discovery of PrEP	Pre-exposure prophylaxis (PrEP) was discovered as a highly effective method for preventing HIV transmission in high-risk populations. This breakthrough transformed HIV prevention strategies and provided an alternative to condom use.
Identification of High-risk Populations	Researchers have used machine learning algorithms to identify subpopulations at higher risk of HIV infection, such as men who have sex with men, transgender women, and people who inject drugs. This knowledge has helped to target HIV prevention efforts more effectively.
Development of Early Warning Systems	Machine learning models have been developed to predict HIV outbreaks in real-time by analyzing trends in social media data, Google searches, and other non-traditional sources. This has the potential to allow for more rapid responses to emerging outbreaks.
Improved Accuracy of Predictive Models	Advances in machine learning algorithms and data analytics have resulted in more accurate predictive models for PrEP and HIV transmission risk. These models can incorporate complex data sources, such as social network analysis, to improve their accuracy.
Development of Tailored Interventions	Machine learning models have been used to identify individual-level risk factors for HIV transmission and PrEP adherence, allowing for tailored interventions to be developed for each patient. This personalized approach has the potential to improve the effectiveness of HIV prevention efforts.

2. Machine learning for PrEP and HIV prediction

Machine learning [5] has been found to be useful both in HIV risk prediction and as a decision support tool for guiding pre-exposure prophylaxis (PrEP) treatment. More recent studies have developed machine learning algorithms using electronic health records data to identify patients at risk of acquiring HIV, hence resulting to discussion of PrEP with the health provider [6]. These electronic health records contain rich data for HIV risk prediction by using the following characteristics: demographic features, social history, diagnosis, laboratory tests and results and the prescriptions used.

In [7], machine learning has been defined as the process by which computational and statistical algorithms "learn" from data, usually with limited human input. This has helped in formation of algorithms that can increase their performance through prediction, pattern recognition, classification and regression using the provided data. Machine learning algorithms range in their complexity [8], [9], [10], [11]. Traditional regression models are used for basic machine learning algorithms. They classify dataset as input and objective [12].To satisfy modeling goal, these machines learn interdependently and are likely to maximize the probability [13]. In most cases, the focus is on the role of machine learning in HIV prevention and to improve prediction in terms of risk for HIV and PrEP treatment. Being data driven, machine learning algorithms can automatically learn from data that identifies complex, nonlinear patterns, and exploiting complex interactions between risk factors [14].This is the advantage they pose over developing predictive models [15], [16] by not requiring statistical inferences or assumptions. These models have been used in prediction of future risks of other conditions like Alzheimer's disease [17], Myocardial infarction [18], suicide [19] and type 2 diabetes [20]. Authors in [21] used machine learning in prediction of HIV/AIDs patients in Guangxi, China.

The authors in [22] identified that machine learning techniques [23] that can help to identify undiagnosed PLWH with a fairly high level of accuracy. The use of machine learning technique has decreased bias in incomplete discovery of HIV status and HIV RNA levels when assessing population level in terms of viral HIV suppression [24]. Moreover, in observational setting, it has reduced the bias on the oversight of the regression model. This allows flexible control of measured confounders [25], suggesting that fewer features reduce the risk of model over-fitting and leads to improvement of the algorithm [26]-[29].

In addition, machine learning algorithms are fast-expanding research areas which are finding their way in HIV research. In machine learning for prediction, AI has also been used to facilitate HIV sero-disclosure [30]. In a small pilot study in [31], the authors have developed and evaluated the Tough Talks virtual reality program to help young MSM role-play HIV serostatus disclosure, with the goal of increasing protective behaviors against HIV transmission. The authors gathered qualitative data through focus groups with young MSM living with HIV on their HIV sero-disclosure experiences. These were then used to create a database of utterances that commonly occur in discussions about HIV serostatus. Participants could pick a virtual character (i.e., an avatar) and role-play various disclosure scenarios with the virtual reality program. Some participants found the tool acceptable but some found it unnecessarily complex and cumbersome to use. A randomized trial is planned to test the effect of this AI tool, delivered online versus in a clinic setting, on HIV viral load and condomless anal sex among young MSM [32].

Artificial intelligence has been introduced into the healthcare field as a means of improving the exactness and accuracy while reducing the number of time-consuming tasks that require human intervention [7], [33], [34], [35]. Because of its ease of use, this innovation could provide a useful tool, allowing for quicker intervention [36]. Various machine learning algorithms have been used in clinical datasets and biomedical which have become the area of interest in medical field [37]-[40]. This might help in controlling the disadvantages of analytical approaches used currently in risk prediction. Here, larger datasets are fed in computer algorithms having many multidimensional variables with high -dimensional and nonlinear relationships among clinical features which can make predictions [41] that are data-driven. Machine learning applications have been extensively utilized on clinical features for cancer and tumor prognosis prediction, in lung cancer and breast cancer. The authors in [42] suggested that Logistic Regression (LR), Gaussian Naive Bayes (GNB), Decision Trees (DT), K-nearest Neighbor (KNN), extreme Gradient Boosting (XGB), Random Forest (RF), and others when used on data from clinical care of HIV patients, they showed to be effective in modeling viral load and CD4 related outcomes. In addition, Xtreme Gradient Boost machine learning [43]-[47] has shown to be effective in prediction of the hospitalization outcome of HIV/AIDs patients with marneffei infection. This is through the prediction of mortality and high-risk factor found in the *talaromycosis* population [21].

The authors in [48] employed five standard ML algorithms, K nearest neighbor, Decision Tree, GNB, Support Vector Machine (SVM), LR, and three ensemble algorithms Gradient Boosting (GB), XGB, and RF to predict the viral load and CD4 status of adults living with HIV/AIDS and enrolled in ART care in Ethiopia. Their results showed that extreme gradient boost and random forest machine learning algorithms performed better when tested with other machine learning models in prediction of viral suppression of individuals enrolled on ART. This showed important connections of supervised machine learning in clinical setting. The researchers in [49] used the following algorithms for HIV- related risk behaviors: Support Vector Machines (SVM); Logistic regression (LG); Decision Tree (DT); and, Random forests (RF). RF indicated good generalizability with future similar samples by proving to be of high predictive accuracy after testing data.So far it has been used to predict progression of HIV and ART optimization [50]-[52]. Likewise, researchers in [53] have used machine learning in developing a multiscale modelling [54] of the HIV-1 infection in the presence of NRTI therapy.

The U.S preventive services Task force in 2019 issued a grade A recommendation for use of PrEP in people with risk of acquiring HIV, noting on improvement of tools to identify potential PrEP candidates. Moreover, several HIV risk scores have been identified for women in sub-Sahara Africa [55], [56], [57]. This proved that machine learning improved the competence and credibility of HIV risk classification, contrary to risk group and model-based approaches, among both men and women and also younger ones (aged 15-24) and older adults. In Eastern Africa, machine learning has been used to identify potential candidates for PrEP counseling via an inclusive approach based on sero-different partnerships [58]-[60] in epidemic settings of Uganda and Kenya [61]. Researchers in [62] used machine learning [63] in risk score identification. ML algorithms employed on data from routine clinical care of HIV patients such as Logistic Regression (LR) [64], Gaussian Naive Bayes (GNB) [65], Decision Trees (DT) [66], K-nearest Neighbor (KNN) [67], eXtreme Gradient Boosting (XGB) [68], Random Forest (RF), and others were effective to model viral load and CD4-related outcomes [42].

Researchers in [69] used machine learning and created a viable machine learning model [70] using a digital survey data with interlinking potential utility in directing health resources including PrEP towards area of great potential gain. This is by using a guided tool for assessing risk behavior to contacting HIV in South Africa. In addition, researchers in [71] used machine learning as a potential utility in a decision making on PrEP by predicting HIV/AIDS knowledge among adolescents and young adult population in Peru. This is through the identification of individuals at high risk of HIV and low conception on HIV. In recent years, the use of predictive models [72]-[76] has helped in the study of prevention, diagnosis and treatment of the HIV/AIDS epidemic. This is facilitated by coming up with new aspects that would be more efficient in treatment and management of the epidemic.

Machine learning can exhibit fairly high level of precision in identifying undiagnosed people living with HIV by learning from nation-wide electronic registry data [77]-[81]. These algorithms may assist in the identification of PrEP candidates and making decision around PrEP [82]. Moreover, they have shown to be very effective in identifying and prediction [83] of HIV risk across both low-income and high-income setting [84]. This has led to facilitation of implementation of preventive strategies like Pre-exposure prophylaxis (PrEP) which has shown its effectiveness in prevention of transmission of HIV by 100% [85]-[88].

Sometimes ethical issues arise when using individual characteristics for clinical prediction. This is because patients may need to consent to using their personal data [89]. Therefore, considerations should be taken to address issues and check if the benefits outweigh the risk of harm to the patient. The researchers in [90] and [91] found out that the risk prediction tools in HIV based on Centers for Disease Control and the prevention criteria for HIV use, in sexual behaviors and STIs, has underated HIV risk among black men who have sex with men. In addition, the authors in [92] assessed the performance of the algorithm by race in the Kaiser Permanente study. The results showed the ability to predict [93] HIV acquisition among black and white patients. However, most of the application only used variables related to sexual orientation and STIs had lower sensitivity for black compared with white patients. This observation is in agreement with the studies carried out in [94]-[96].

Table 2 provides a brief overview of how various machine learning algorithms have been used in PrEP and HIV prediction. The purpose of each algorithm, the data source used, and the results achieved are presented.

Table 2 Machine learning for PrEP and HIV prediction

Algorithm	**Purpose**	**Data Source**	**Results**
Logistic Regression	Predicting HIV Risk	Electronic Health Records	Achieved an AUC of 0.73 for predicting HIV infection within the next year in a cohort of MSM.
Random Forest	Identifying PrEP Candidates	Demographic and Behavioral Data	Achieved an AUC of 0.79 for identifying individuals who could benefit from PrEP in a cohort of high-risk individuals.
Deep Neural Networks	Predicting HIV Infection	Clinical and Behavioral Data	Achieved an AUC of 0.85 for predicting HIV infection within the next year in a cohort of high-risk individuals.
Support Vector Machines	Identifying HIV-Infected Individuals	Laboratory and Clinical Data	Achieved a sensitivity of 0.86 and specificity of 0.93 for identifying HIV-infected individuals in a cohort of patients seeking HIV testing.
Decision Trees	Predicting PrEP Adherence	Self-Reported Adherence Data	Achieved an accuracy of 70% for predicting PrEP adherence in a cohort of MSM.

3. Research gaps

While machine learning has shown promise in the field of PrEP and HIV prediction, there are still several research gaps that need to be addressed to improve the development of machine learning models in this area. Some of these gaps include:

- Lack of standardized datasets: Machine learning models require large, standardized datasets to train and validate their accuracy. However, there is a lack of standardized datasets in the field of PrEP and HIV prediction, which limits the development and comparison of machine learning models.
- Need for more diverse data sources: Current datasets used for machine learning models in PrEP and HIV prediction are often limited to clinical data or self-reported data from individuals. There is a need for more diverse data sources, such as data from social media, wearable devices, or electronic health records, to improve the accuracy and inclusivity of machine learning models.
- Limited generalizability of models: Machine learning models developed for PrEP and HIV prediction are often trained on data from specific populations or geographic regions, which limits their generalizability to other populations or regions.
- Ethical concerns: The development of machine learning models for PrEP and HIV prediction raises ethical concerns regarding privacy, data security, and potential algorithmic bias. Further research is needed to address these concerns and ensure the ethical use of machine learning in this area.
- Lack of interpretability: Many machine learning models are considered "black boxes," meaning that it is difficult to understand how the model arrived at its predictions. There is a need for more interpretable machine learning models that can help healthcare providers and researchers understand the factors contributing to an individual's risk of HIV infection.

In addition, developing a model for PrEP and HIV can be challenging due to various reasons including:

- Complexity of the virus: HIV is a complex virus that mutates rapidly and has multiple strains. Developing a model that accurately predicts HIV infection and PrEP efficacy requires a deep understanding of the virus's behavior and evolution.
- Lack of data: Developing an accurate model for PrEP and HIV requires large amount of data from diverse populations. However, there are still many gaps in data collection, particularly in low-income countries and among marginalized communities.
- Ethical considerations: Developing these models requires careful consideration of ethical issues, including privacy, confidentiality and informed consent. There is need to balance potential benefits of the model against the risk of unintended harm, such as stigmatization or discrimination.

Table 3 highlights some of the challenges associated with the use of machine learning algorithms in the context of PrEP and HIV prediction. The specific challenges may vary depending on the study or dataset. Additionally, some challenges may be interdependent and exacerbate each other.

Table 3 Machine learning-based PrEP and HIV prediction challenges

Challenge	Description
Lack of Standardized Data	Machine learning models require large and standardized datasets to achieve optimal accuracy. However, in the context of PrEP and HIV prediction, there is a lack of standardized data, which limits the ability to develop and compare machine learning models.
Limited Generalizability	Machine learning models developed for PrEP and HIV prediction are often trained on data from specific populations or geographic regions, which limits their generalizability to other populations or regions.
Data Bias and Privacy Concerns	Machine learning algorithms can be influenced by bias in the data used to train them, leading to discriminatory outcomes. Moreover, privacy concerns may arise due to the use of sensitive information in developing machine learning models.
Limited Interpretability	Many machine learning models are considered "black boxes," which makes it difficult to understand how they arrive at their predictions. This lack of interpretability can make it difficult for clinicians and researchers to use the models effectively.
Data Imbalance	HIV is a rare event, which can lead to imbalanced datasets, where there are significantly fewer positive cases than negative cases. This can result in machine learning models with poor predictive performance.
Lack of Expertise	Developing and implementing machine learning models requires a high level of expertise in both data science and clinical domains. The shortage of such expertise in the healthcare sector can be a significant challenge.

Addressing these research gaps will be critical in advancing the development of machine learning models for PrEP and HIV prediction and improving their accuracy and impact on HIV prevention efforts.

4. Conclusion

In recent years, machine learning has shown promising results in the field of PrEP (Pre-exposure prophylaxis) and HIV prediction. Machine learning algorithms have the potential to improve the accuracy and speed of predicting individuals who are at risk of HIV infection and could benefit from PrEP. By analyzing a wide range of variables such as demographic data, behavioral patterns, and clinical markers, machine learning models can identify patterns and correlations that may not be apparent to human experts. These models can also continuously learn from new data, improving their predictive capabilities over time. The use of machine learning for PrEP and HIV prediction could have a significant impact on reducing the spread of HIV and increasing access to PrEP for those who need it the most. However, there are also potential challenges and ethical concerns to consider, such as the potential for algorithmic bias and the need for privacy and data security. In conclusion, while the use of machine learning for PrEP and HIV prediction is still in its early stages, the results so far are promising. Further research and development in this area have the potential to improve the accuracy and efficiency of HIV prevention efforts and ultimately help to curb the spread of the virus.

Compliance with ethical standards

Acknowledgments

Special appreciation goes to all my colleagues who supported me when writing this manuscript.

References

[1] Ajayi AI, Awopegba OE, Adeagbo OA, Ushie BA. Low coverage of HIV testing among adolescents and young adults in Nigeria: Implication for achieving the UNAIDS first 95. PloS one. 2020 May 19, 15(5):e0233368.

[2] Frank TD, Carter A, Jahagirdar D, Biehl MH, Douwes-Schultz D, Larson SL, Arora M, Dwyer-Lindgren L, Steuben KM, Abbastabar H, Abu-Raddad LJ. Global, regional, and national incidence, prevalence, and mortality of HIV, 1980–2017, and forecasts to 2030, for 195 countries and territories: a systematic analysis for the Global Burden of Diseases, Injuries, and Risk Factors Study 2017. The lancet HIV. 2019 Dec 1, 6(12):e831-59.

[3] Pandey A, Galvani AP. The global burden of HIV and prospects for control. The Lancet HIV. 2019 Dec 1, 6(12):e809-11.

[4] Fieggen J, Smith E, Arora L, Segal B. The role of machine learning in HIV risk prediction. Frontiers in Reproductive Health. 2022, 4.

[5] Honi DG, Ali AH, Abduljabbar ZA, Ma J, Nyangaresi VO, Mutlaq KA, Umran SM. Towards Fast Edge Detection Approach for Industrial Products. In2022 IEEE 21st International Conference on Ubiquitous Computing and Communications (IUCC/CIT/DSCI/SmartCNS) 2022 Dec 19 (pp. 239-244). IEEE.

[6] Ortblad KF, Baeten JM. Electronic health record tools to catalyse PrEP conversations. The Lancet HIV. 2019 Oct 1, 6(10):e644-5.

[7] Marcus JL, Hurley LB, Krakower DS, Alexeeff S, Silverberg MJ, Volk JE. Use of electronic health record data and machine learning to identify candidates for HIV pre-exposure prophylaxis: a modelling study. The lancet HIV. 2019 Oct 1, 6(10):e688-95.

[8] Singh A, Thakur N, Sharma A. A review of supervised machine learning algorithms. In2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom) 2016 Mar 16 (pp. 1310-1315). Ieee.

[9] Wang S, Jia D, Weng X. Deep reinforcement learning for autonomous driving. arXiv preprint arXiv:1811.11329. 2018 Nov 28.

[10] Anand A, Haque MA, Alex JS, Venkatesan N. Evaluation of Machine learning and Deep learning algorithms combined with dimentionality reduction techniques for classification of Parkinson's Disease. In2018 IEEE international symposium on signal processing and information technology (ISSPIT) 2018 Dec 6 (pp. 342-347). IEEE.

[11] Nyangaresi VO, Rodrigues AJ. Efficient handover protocol for 5G and beyond networks. Computers & Security. 2022 Feb 1, 113:102546.

[12] Maulud D, Abdulazeez AM. A review on linear regression comprehensive in machine learning. Journal of Applied Science and Technology Trends. 2020 Dec 31, 1(4):140-7.

[13] Mahesh B. Machine learning algorithms-a review. International Journal of Science and Research (IJSR).[Internet]. 2020 Jan, 9:381-6.

[14] Patel B, Sengupta P. Machine learning for predicting cardiac events: what does the future hold?. Expert review of cardiovascular therapy. 2020 Feb 1, 18(2):77-84.

[15] Waljee AK, Higgins PD, Singal AG. A primer on predictive models. Clinical and translational gastroenterology. 2014 Jan, 5(1):e44.

[16] Nyangaresi VO, Abduljabbar ZA, Al Sibahee MA, Ibrahim A, Yahya AN, Abduljaleel IQ, Abood EW. Optimized Hysteresis Region Authenticated Handover for 5G HetNets. InArtificial Intelligence and Sustainable Computing: Proceedings of ICSISCET 2021 2022 Nov 16 (pp. 91-111). Singapore: Springer Nature Singapore.

[17] Park JH, Cho HE, Kim JH, Wall MM, Stern Y, Lim H, Yoo S, Kim HS, Cha J. Machine learning prediction of incidence of Alzheimer's disease using large-scale administrative health data. NPJ digital medicine. 2020 Mar 26, 3(1):46.

[18] Kwiecinski J, Tzolos E, Meah MN, Cadet S, Adamson PD, Grodecki K, Joshi NV, Moss AJ, Williams MC, van Beek EJ, Berman DS. Machine learning with 18F-sodium fluoride PET and quantitative plaque analysis on CT angiography for the future risk of myocardial infarction. Journal of Nuclear Medicine. 2022 Jan 1, 63(1):158-65.

[19] Roy A, Nikolitch K, McGinn R, Jinah S, Klement W, Kaminsky ZA. A machine learning approach predicts future risk to suicidal ideation from social media data. NPJ digital medicine. 2020 May 26, 3(1):78.

[20] Farran B, AlWotayan R, Alkandari H, Al-Abdulrazzaq D, Channanath A, Thanaraj TA. Use of non-invasive parameters and machine-learning algorithms for predicting future risk of type 2 diabetes: a retrospective cohort study of health data from Kuwait. Frontiers in endocrinology. 2019 Sep 11, 10:624.

[21] Shi M, Lin J, Wei W, Qin Y, Meng S, Chen X, Li Y, Chen R, Yuan Z, Qin Y, Huang J. Machine learning-based in-hospital mortality prediction of HIV/AIDS patients with Talaromyces marneffei infection in Guangxi, China. PLOS Neglected Tropical Diseases. 2022 May 4, 16(5):e0010388.

[22] Ahlström MG, Ronit A, Omland LH, Vedel S, Obel N. Algorithmic prediction of HIV status using nation-wide electronic registry data. EClinicalMedicine. 2019 Dec 1, 17:100203.

[23] Nyangaresi VO, El-Omari NK, Nyakina JN. Efficient Feature Selection and ML Algorithm for Accurate Diagnostics. Journal of Computer Science Research. 2022 Jan 25, 4(1):10-9.

[24] Balzer LB, Ayieko J, Kwarisiima D, Chamie G, Charlebois ED, Schwab J, van der Laan MJ, Kamya MR, Havlir DV, Petersen ML. Far from MCAR: obtaining population-level estimates of HIV viral suppression. Epidemiology (Cambridge, Mass.). 2020 Sep, 31(5):620.

[25] Tran L, Yiannoutsos CT, Musick BS, Wools-Kaloustian KK, Siika A, Kimaiyo S, van der Laan MJ, Petersen M. Evaluating the impact of a HIV low-risk express care task-shifting program: a case study of the targeted learning roadmap. Epidemiologic methods. 2016 Dec 1, 5(1):69-91.

[26] Peng Y, Nagata MH. An empirical overview of nonlinearity and overfitting in machine learning using COVID-19 data. Chaos, Solitons & Fractals. 2020 Oct 1, 139:110055.

[27] Ying X. An overview of overfitting and its solutions. InJournal of physics: Conference series 2019 Feb (Vol. 1168, p. 022022). IOP Publishing.

[28] Yeom S, Giacomelli I, Fredrikson M, Jha S. Privacy risk in machine learning: Analyzing the connection to overfitting. In2018 IEEE 31st computer security foundations symposium (CSF) 2018 Jul 9 (pp. 268-282). IEEE.

[29] Nyangaresi VO, Rodrigues AJ, Abeka SO. ANN-FL secure handover protocol for 5G and beyond networks. InTowards new e-Infrastructure and e-Services for Developing Countries: 12th EAI International Conference, AFRICOMM 2020, Ebène City, Mauritius, December 2-4, 2020, Proceedings 12 2021 (pp. 99-118). Springer International Publishing.

[30] Marcus JL, Sewell WC, Balzer LB, Krakower DS. Artificial intelligence and machine learning for HIV prevention: emerging approaches to ending the epidemic. Current HIV/AIDS Reports. 2020 Jun, 17:171-9.

[31] Muessig KE, Knudtson KA, Soni K, Larsen MA, Traum D, Dong W, Conserve DF, Leuski A, Artstein R, Hightow-Weidman LB. “I didn’t tell you sooner because i didn’t know how to handle it myself.” Developing a virtual reality program to support HIV-status disclosure decisions. Digital culture & education. 2018, 10:22.

[32] Makhema J, Wirth KE, Pretorius Holme M, Gaolathe T, Mmalane M, Kadima E, Chakalisa U, Bennett K, Leidner J, Manyake K, Mbikiwa AM. Universal testing, expanded treatment, and incidence of HIV infection in Botswana. New England Journal of Medicine. 2019 Jul 18, 381(3):230-42.

[33] Yu KH, Beam AL, Kohane IS. Artificial intelligence in healthcare. Nature biomedical engineering. 2018 Oct, 2(10):719-31.

[34] Jiang F, Jiang Y, Zhi H, Dong Y, Li H, Ma S, Wang Y, Dong Q, Shen H, Wang Y. Artificial intelligence in healthcare: past, present and future. Stroke and vascular neurology. 2017 Dec 1, 2(4).

[35] Ghrabat MJ, Hussien ZA, Khalefa MS, Abduljabba ZA, Nyangaresi VO, Al Sibahee MA, Abood EW. Fully automated model on breast cancer classification using deep learning classifiers. Indonesian Journal of Electrical Engineering and Computer Science. 2022 Oct, 28(1):183-91.

[36] Rong G, Mendez A, Assi EB, Zhao B, Sawan M. Artificial intelligence in healthcare: review and prediction case studies. Engineering. 2020 Mar 1, 6(3):291-301.

[37] Jayatilake SM, Ganegoda GU. Involvement of machine learning tools in healthcare decision making. Journal of healthcare engineering. 2021 Jan 27, 2021.

[38] Waring J, Lindvall C, Umeton R. Automated machine learning: Review of the state-of-the-art and opportunities for healthcare. Artificial intelligence in medicine. 2020 Apr 1, 104:101822.

[39] Haque IR, Neubert J. Deep learning approaches to biomedical image segmentation. Informatics in Medicine Unlocked. 2020 Jan 1, 18:100297.

[40] Solares JR, Raimondi FE, Zhu Y, Rahimian F, Canoy D, Tran J, Gomes AC, Payberah AH, Zottoli M, Nazarzadeh M, Conrad N. Deep learning for electronic health records: A comparative review of multiple deep neural architectures. Journal of biomedical informatics. 2020 Jan 1, 101:103337.

[41] Nyangaresi VO, Ahmad M, Alkhayyat A, Feng W. Artificial neural network and symmetric key cryptography based verification protocol for 5G enabled Internet of Things. Expert Systems. 2022 Dec, 39(10):e13126.

[42] Romero-Rodríguez E, Pérula-de Torres LA, González-Lama J, Jiménez-García C, Castro-Jiménez RA, González-Bernal JJ, Rodríguez-Fernández P, Mielgo-Ayuso J, Santamaría-Peláez M, González-Santos J. Clinical Presentation of the SARS-CoV-2 Virus Infection and Predictive Validity of the PCR Test in Primary Health Care Worker Patients of the Spanish National Health System. Journal of Clinical Medicine. 2022 Jan 4, 11(1):243.

[43] Wang F, Ross CL. Machine learning travel mode choices: Comparing the performance of an extreme gradient boosting model with a multinomial logit model. Transportation Research Record. 2018 Dec, 2672(47):35-45.

[44] Pathy A, Meher S, Balasubramanian P. Predicting algal biochar yield using eXtreme Gradient Boosting (XGB) algorithm of machine learning methods. Algal Research. 2020 Sep 1, 50:102006.

[45] Babajide Mustapha I, Saeed F. Bioactive molecule prediction using extreme gradient boosting. Molecules. 2016 Jul 28, 21(8):983.

[46] Yang M, Tao B, Chen C, Jia W, Sun S, Zhang T, Wang X. Machine learning models based on molecular fingerprints and an extreme gradient boosting method lead to the discovery of JAK2 inhibitors. Journal of Chemical Information and Modeling. 2019 Nov 20, 59(12):5002-12.

[47] Al Sibahee MA, Ma J, Nyangaresi VO, Abduljabbar ZA. Efficient Extreme Gradient Boosting Based Algorithm for QoS Optimization in Inter-Radio Access Technology Handoffs. In2022 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA) 2022 Jun 9 (pp. 1-6). IEEE.

[48] Seboka BT, Yehualashet DE, Tesfa GA. Artificial Intelligence and Machine Learning Based Prediction of Viral Load and CD4 Status of People Living with HIV (PLWH) on Anti-Retroviral Treatment in Gedeo Zone Public Hospitals. International Journal of General Medicine. 2023 Dec 31:435-51.

[49] Wang B, Liu F, Deveaux L, Ash A, Gerber B, Allison J, Herbert C, Poitier M, MacDonell K, Li X, Stanton B. Predicting Adolescent Intervention Non-responsiveness for Precision HIV Prevention Using Machine Learning. AIDS and Behavior. 2022 Oct 18:1-1.

[50] Singh Y. Machine learning to improve the effectiveness of ANRS in predicting HIV drug resistance. Healthcare informatics research. 2017 Oct 31, 23(4):271-6.

[51] Riemenschneider M, Heider D. Current approaches in computational drug resistance prediction in HIV. Current HIV research. 2016 Jul 1, 14(4):307-15.

[52] Zazzi M, Cozzi-Lepri A, Prosperi MC. Computer-aided optimization of combined anti-retroviral therapy for HIV: new drugs, new drug targets and drug resistance. Current HIV research. 2016 Mar 1, 14(2):101-9.

[53] Tunc H, Sari M, Kotil S. Machine learning aided multiscale modelling of the HIV-1 infection in the presence of NRTI therapy. PeerJ. 2023 Mar 31, 11:e15033.

[54] Nyangaresi VO, Rodrigues AJ, Abeka SO. Machine learning protocol for secure 5G handovers. International Journal of Wireless Information Networks. 2022 Mar, 29(1):14-35.

[55] Balkus JE, Brown E, Palanee T, Nair G, Gafoor Z, Zhang J, Richardson BA, Chirenje ZM, Marrazzo JM, Baeten JM. An empiric HIV risk scoring tool to predict HIV-1 acquisition in African women. Journal of acquired immune deficiency syndromes (1999). 2016 Jul 7, 72(3):333.

[56] Pintye J, Drake AL, Kinuthia J, Unger JA, Matemo D, Heffron RA, Barnabas RV, Kohler P, McClelland RS, John-Stewart G. A risk assessment tool for identifying pregnant and postpartum women who may benefit from preexposure prophylaxis. Clinical Infectious Diseases. 2017 Mar 15, 64(6):751-8.

[57] Balzer LB, Havlir DV, Kamya MR, Chamie G, Charlebois ED, Clark TD, Koss CA, Kwarisiima D, Ayieko J, Sang N, Kabami J. Machine learning to identify persons at high-risk of human immunodeficiency virus acquisition in rural Kenya and Uganda. Clinical Infectious Diseases. 2020 Nov 1, 71(9):2326-33.

[58] Reis RK, Melo ES, Fernandes NM, Antonini M, Neves LA, Gir E. Inconsistent condom use between serodifferent sexual partnerships to the human immunodeficiency virus. Revistalatino-americana de enfermagem. 2019 Dec 5, 27.

[59] Rodger AJ, Cambiano V, Bruun T, Vernazza P, Collins S, Degen O, Corbelli GM, Estrada V, Geretti AM, Beloukas A, Raben D. Risk of HIV transmission through condomless sex in serodifferent gay couples with the HIV-positive

partner taking suppressive antiretroviral therapy (PARTNER): final results of a multicentre, prospective, observational study. The Lancet. 2019 Jun 15, 393(10189):2428-38.

[60] Baral S, Rao A, Sullivan P, Phaswana-Mafuya N, Diouf D, Millett G, Musyoki H, Geng E, Mishra S. The disconnect between individual-level and population-level HIV prevention benefits of antiretroviral treatment. The lancet HIV. 2019 Sep 1, 6(9):e632-8.

[61] Havlir DV, Balzer LB, Charlebois ED, Clark TD, Kwarisiima D, Ayieko J, Kabami J, Sang N, Liegler T, Chamie G, Camlin CS. HIV testing and treatment with the use of a community health approach in rural Africa. New England Journal of Medicine. 2019 Jul 18, 381(3):219-29.

[62] Zheng W, Balzer L, van der Laan M, Petersen M, SEARCH Collaboration. Constrained binary classification using ensemble learning: an application to cost-efficient targeted PrEP strategies. Statistics in medicine. 2018 Jan 30, 37(2):261-79.

[63] Nyangaresi VO, Rodrigues AJ, Abeka SO. Neuro-fuzzy based handover authentication protocol for ultra dense 5G networks. In2020 2nd Global Power, Energy and Communication Conference (GPECOM) 2020 Oct 20 (pp. 339-344). IEEE.

[64] Nwazelibe VE, Unigwe CO, Egbueri JC. Integration and comparison of algorithmic weight of evidence and logistic regression in landslide susceptibility mapping of the orumba north erosion-prone region, nigeria. Modeling Earth Systems and Environment. 2023 Mar, 9(1):967-86.

[65] Li J, Li Z, Wen C, Peng Q, Huang P. Train operation conflict detection for high-speed railways: a naïve Bayes approach. International Journal of Rail Transportation. 2023 Mar 4, 11(2):188-206.

[66] Douiba M, Benkirane S, Guezzaz A, Azrour M. An improved anomaly detection model for IoT security using decision tree and gradient boosting. The Journal of Supercomputing. 2023 Feb, 79(3):3392-411.

[67] Wang AX, Chukova SS, Nguyen BP. Ensemble k-nearest neighbors based on centroid displacement. Information Sciences. 2023 Jun 1, 629:313-23.

[68] Zhang W, He Y, Wang L, Liu S, Meng X. Landslide Susceptibility mapping using random forest and extreme gradient boosting: A case study of Fengjie, Chongqing. Geological Journal. 2023 Feb 7.

[69] Majam M, Phatsoane M, Hanna K, Faul C, Arora L, Makthal S, Kumar A, Jois K, Lalla-Edward ST. Utility of a machine-guided tool for assessing risk behavior associated with contracting HIV in three sites in South Africa: protocol for an in-field evaluation. JMIR Research Protocols. 2021 Dec 2, 10(12):e30304.

[70] Nyangaresi VO, Abeka SO, Rodrigues AJ. Delay sensitive protocol for high availability LTE handovers. American Journal of Networks and Communications. 2020 Feb, 9(1): 1-10.

[71] Aybar-Flores A, Talavera A, Espinoza-Portilla E. Predicting the HIV/AIDS Knowledge among the Adolescent and Young Adult Population in Peru: Application of Quasi-Binomial Logistic Regression and Machine Learning Algorithms. International Journal of Environmental Research and Public Health. 2023 Mar 30, 20(7):5318.

[72] Wathore R, Rawlekar S, Anjum S, Gupta A, Bherwani H, Labhasetwar N, Kumar R. Improving performance of deep learning predictive models for COVID-19 by incorporating environmental parameters. Gondwana Research. 2023 Feb 1, 114:69-77.

[73] Aldrees A, Javed MF, Taha AT, Mohamed AM, Jasiński M, Gono M. Evolutionary and ensemble machine learning predictive models for evaluation of water quality. Journal of Hydrology: Regional Studies. 2023 Apr 1, 46:101331.

[74] Huang HH, Hsieh SJ, Chen MS, Jhou MJ, Liu TC, Shen HL, Yang CT, Hung CC, Yu YY, Lu CJ. Machine Learning Predictive Models for Evaluating Risk Factors Affecting Sperm Count: Predictions Based on Health Screening Indicators. Journal of Clinical Medicine. 2023 Feb 3, 12(3):1220.

[75] Li Q, Lv H, Chen Y, Shen J, Shi J, Zhou C. Development and Validation of a Machine Learning Predictive Model for Cardiac Surgery-Associated Acute Kidney Injury. Journal of Clinical Medicine. 2023 Feb 1, 12(3):1166.

[76] Nyangaresi VO, Abeka SO, Rodgrigues A. Secure timing advance based context-aware handover protocol for vehicular ad-hoc heterogeneous networks. International Journal of Cyber-Security and Digital Forensics. 2018 Sep 1, 7(3):256-75.

[77] Ahlström MG, Ronit A, Omland LH, Vedel S, Obel N. Algorithmic prediction of HIV status using nation-wide electronic registry data. EClinicalMedicine. 2019 Dec 1, 17:100203.

[78] Alehegn M. Application of machine learning and deep learning for the prediction of HIV/AIDS. HIV & AIDS Review. International Journal of HIV-Related Problems. 2022, 21(1):17-23.

[79] Qiao S, Li X, Olatosi B, Young SD. Utilizing Big Data analytics and electronic health record data in HIV prevention, treatment, and care research: a literature review. AIDS care. 2021 Jul 7:1-21.

[80] Bao Y, Medland NA, Fairley CK, Wu J, Shang X, Chow EP, 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. Journal of Infection. 2021 Jan 1, 82(1):48-59.

[81] Orel E, Esra R, Estill J, Marchand-Maillet S, Merzouki A, Keiser O. Machine learning to identify socio-behavioural predictors of HIV positivity in East and Southern Africa. medRxiv. 2020 Jan 27:2020-01.

[82] Krakower DS, Gruber S, Hsu K, Menchaca JT, Maro JC, Kruskal BA, Wilson IB, Mayer KH, Klompas M. Development and validation of an automated HIV prediction algorithm to identify candidates for pre-exposure prophylaxis: a modelling study. The Lancet HIV. 2019 Oct 1, 6(10):e696-704.

[83] Nyangaresi VO. Target Tracking Area Selection and Handover Security in Cellular Networks: A Machine Learning Approach. InProceedings of Third International Conference on Sustainable Expert Systems: ICSES 2022 2023 Feb 23 (pp. 797-816). Singapore: Springer Nature Singapore.

[84] Xu X, Ge Z, Chow EP, Yu Z, Lee D, Wu J, Ong JJ, Fairley CK, Zhang L. A machine-learning-based risk-prediction tool for HIV and sexually transmitted infections acquisition over the next 12 months. Journal of Clinical Medicine. 2022 Mar 25, 11(7):1818.

[85] McGillen JB, Anderson SJ, Dybul MR, Hallett TB. Optimum resource allocation to reduce HIV incidence across sub-Saharan Africa: a mathematical modelling study. The lancet HIV. 2016 Sep 1, 3(9):e441-8.

[86] Rodger AJ, Cambiano V, Bruun T, Vernazza P, Collins S, Degen O, Corbelli GM, Estrada V, Geretti AM, Beloukas A, Raben D. Risk of HIV transmission through condomless sex in serodifferent gay couples with the HIV-positive partner taking suppressive antiretroviral therapy (PARTNER): final results of a multicentre, prospective, observational study. The Lancet. 2019 Jun 15, 393(10189):2428-38.

[87] McCormack S, Dunn DT, Desai M, Dolling DI, Gafos M, Gilson R, Sullivan AK, Clarke A, Reeves I, Schembri G, Mackie N. Pre-exposure prophylaxis to prevent the acquisition of HIV-1 infection (PROUD): effectiveness results from the pilot phase of a pragmatic open-label randomised trial. The Lancet. 2016 Jan 2, 387(10013):53-60.

[88] Grant RM, Anderson PL, McMahan V, Liu A, Amico KR, Mehrotra M, Hosek S, Mosquera C, Casapia M, Montoya O, Buchbinder S. An observational study of preexposure prophylaxis uptake, sexual practices, and HIV incidence among men and transgender women who have sex with men. The Lancet. Infectious diseases. 2014 Sep, 14(9):820.

[89] Cato KD, Bockting W, Larson E. Did I tell you that? Ethical issues related to using computational methods to discover non-disclosed patient characteristics. Journal of Empirical Research on Human Research Ethics. 2016 Jul, 11(3):214-9.

[90] Lancki N, Almirol E, Alon L, McNulty M, Schneider JA. PrEP guidelines have low sensitivity for identifying seroconverters in a sample of Young Black men who have sex with men in Chicago. AIDS (London, England). 2018 Jan 1, 32(3):383.

[91] Jones J, Hoenigl M, Siegler AJ, Sullivan PS, Little S, Rosenberg E. Assessing the Performance of 3 Human Immunodeficiency Virus Incidence Risk Scores in a Cohort of Black and White Men Who Have Sex With Men in the South. Sex Transm Dis. 2017 Jan, 44(5):297–302.

[92] Marcus JL, Sewell WC, Balzer LB, Krakower DS. Artificial intelligence and machine learning for HIV prevention: emerging approaches to ending the epidemic. Current HIV/AIDS Reports. 2020 Jun, 17:171-9.

[93] Nyangaresi VO, Rodrigues AJ, Abeka SO. Secure Handover Protocol for High Speed 5GNetworks. International Journal of Advanced Networking and Applications. 2020Mar, 11(06): 4429-4442.

[94] Petroll AE, Mosack KE. Physician awareness of sexual orientation and preventive health recommendations to men who have sex with men. Sexually transmitted diseases. 2011 Jan, 38(1):63.

[95] Estrich CG, Gratzer B, Hotton AL. Differences in Sexual Health, Risk Behaviors, and Substance Use Among Women by Sexual Identity. Sexually Transmitted Diseases. 2014 Mar 1, 41(3):194-9.

[96] Eaton LA, Driffin DD, Smith H, Conway-Washington C, White D, Cherry C. Psychosocial factors related to willingness to use pre-exposure prophylaxis for HIV prevention among Black men who have sex with men attending a community event. Sexual health. 2014 Jul 8, 11(3):244-51.