We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

6,300
Open access books available

170,000
International authors and editors

185M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Chapter

A Deep Learning Approaches for Modeling and Predicting of HIV Test Results Using EDHS Dataset

Daniel Mesafint Belete and Manjaiah D. Huchaiah

Abstract

At present, HIV/AIDS has steadily been listed in the top position as a major cause of death. However, HIV is largely preventable and can be avoided by making strategies to increase HIV early prediction. So, there is a need for a predictive tool that can help the domain experts with early prediction of the disease and hence can recommend strategies to stop the prognosis of the diseases. Using deep learning models, we investigated whether demographic and health survey dataset might be utilized to predict HIV test status. The contribution of this work is to improve the accuracy of a model for predicting an individual's HIV test status. We employed deep learning models to predict HIV status using Ethiopian demography and health survey (EDHS) datasets. Furthermore, we discovered that predictive models based on these dataset may be used to forecast individuals' HIV test status, which might assist domain experts prioritize strategies and policies to safeguard the pandemic. The outcome of the study confirms that a DL model provides the best results with the most promising extracted features. The accuracy of the all DL models can further be enhanced by including the big dataset for predicting the prognosis of the disease.

Keywords: deep learning, prediction, CNNLSTM, CNNRNN, EDHS, HIV/AIDS test result

1. Introduction

HIV is the world's most critical community health and development problem. Although millions of people have died as a result of AIDS since the pandemic began in 1981, around 36 million individuals now have HIV. An estimated 19 million people living with HIV are enrolled in and getting treatment through regular care programs [1].

Local HIV/AIDS epidemics require immediate investigation and development of relevant intervention plans, as well as methodologies. Behavioral and socio-demographic characteristics are key contributors to the spread of HIV and require a study on the nature and the influence of the HIV pandemic in a specific community [2]. Although the fact that HIV testing is an efficient technique for determining the test status of people, even has challenges and limitations. As a result, strong prediction models are critical for managing and monitoring the local HIV pandemic.

Ethiopian demographic and health survey (EDHS-CSA) [3] generates a massive amount of dataset that may be analyzed to extract important evidence. The development of the deep learning (DL) model supports the processing of huge datasets and the extraction of underlying dataset patterns that support decision making.

Deep learning methods have recently achieved noteworthy success in a variety of research disciplines, including speech recognition [4], natural language processing [5], recommendation systems [6], and computer vision [7]. This approach is very useful in the health sector for disease prediction and classification. Deep learning algorithms are one of the most recent breakthroughs in HIV statistical dataset prediction tools and identification approaches, allowing for faster processing of large datasets. These algorithms can also be used to predict disease. These techniques work well and can be used to predict HIV test results.

In this paper, we use numerous deep learning application models to construct an HIV status prediction system. On the Ethiopian demographic and health survey dataset (EDHS), six DL models have been developed and are being deployed. These deep learning models were tested using well-known metrics such as accuracy, precision, recall, AUC, and F1 scores.

Our contributions are presented below concerning the core goal of predicting HIV test status:

- Identification of the best performing deep learning algorithm for the task at hand among well-known and widely utilized ones.

- To conduct an HIV test prediction research using deep learning application models with the EDHS dataset.

- We created the HIV/AIDS dataset by applying various techniques such as dataset acquisition, dataset labeling and compared the findings using cutting-edge methodologies, and we got good findings.

To the best of our understanding, no research has been conducted using deep learning models to predict HIV test status using the EDHS dataset. This is the first time that a deep learning model has been used in the health sector to predict HIV test results using only 20 attributes. This work may motivate researchers to validate models using other HIV/AIDS datasets.

The major goal of this study is to propose the development of a more accurate prognostic tool for HIV/AIDS test result prediction. This research comprises six DL models that were used to conduct detailed analyses on the EDHS dataset. The algorithm comparison is presented in a logical and well-organized manner, allowing DL to produce more effective and prominent findings.

The remaining section is structured as follows. Section 2 discusses relevant research studies. The proposed techniques are presented in Section 3. Section 4 describes an experimental design and results from the analysis. Section 5 is devoted to comparative analysis, while Section 6 is focused on concluding remarks.

2. Related work

Various scholars have previously done a great amount of study on health topics. This section provides an overview of previous research in the prediction of HIV

epidemics using advanced ML algorithms and big dataset technologies. We highlighted some of the most important and significant work done by various researchers in this field.

McSharry et al. [8] used ML approaches to successfully discover HIV predictors through screening on the PHIA dataset. The study aims to analyze the HIV disease trial at various levels of society, detect HIV predictors, and forecast the risk of the disease. For the prediction tasks, six ML models were utilized in the study. The primary finding of this study is that the XGBoost algorithm greatly outperformed the other algorithms in terms of identifying HIV-positive. Another ML methodology presented by Orel et al. [9] examined more than 3,200 parameters of the current Demographic Health Surveys from ten African nations. This study trained four ML models and chose the best one through the f1 score. The primary goal of the study is to identify PLHIV at a rate of more than 95\% and to identify the number of positive persons at a rate greater than 95%. The authors emphasized the significance of attribute extraction strategies in mining information for prediction. Using four separate datasets from UCI, Lu et al. [10] employed one-hot coding to translate the protease cleavage site dataset for prediction using two DL models, RNN and LSTM. Finally, the DL model results are compared to SVM and RF. The author Wang et al. [11] created a convenient model to explain the prevalence of HIV and forecast its occurrence in Guangxi. From 2005 to 2016, the HIV incidence statistics datasets were utilized in the study. They trained the HIV incidence using four models, including LSTM, ARIMA, ES, and GRNN. Following training, all models are assessed using the most popular prediction task evaluation criteria. According to the findings of the studies, LSTM and ARIMA outperform ES and GRNN. The LSTM model, on the other hand, proved more successful than other models. Ahlström et al. [12] provided an algorithmic prediction of HIV status. The author investigated whether a dataset from a national electronic registry might be utilized to predict HIV status using machine learning techniques. The study employed multiple techniques to train prediction models, which were then verified using a dataset from Danish households. They trained the models to simulate various clinical. Steiner et al. [13] evaluate the DL models for drug resistance prediction using the HIV-1 sequencing dataset. DL algorithms are combined with HIV genotypic and phenotypic datasets and studies by the author to study the classification performance of the fundamental evolutionary methods of HIV treatment resistance. They assessed the effectiveness of three DL models using a publicly accessible HIV sequencing dataset.

As a result, we have observed several research projects being conducted in the field of HIV/AIDS prediction. All of the available approaches have been shown to perform on various datasets and produce promising results. These concerns inspire us to investigate deep learning methods for predicting HIV test status to improve prediction performance.

3. Proposed methodology

We discuss the proposed work in this part, which encompasses several phases such as pre-processing, normalizing features, and a deep learning-based prediction technique with parameter settings. **Figure 1** depicts the architecture for the proposed deep learning models for predicting HIV test status in people using the EDHS dataset.
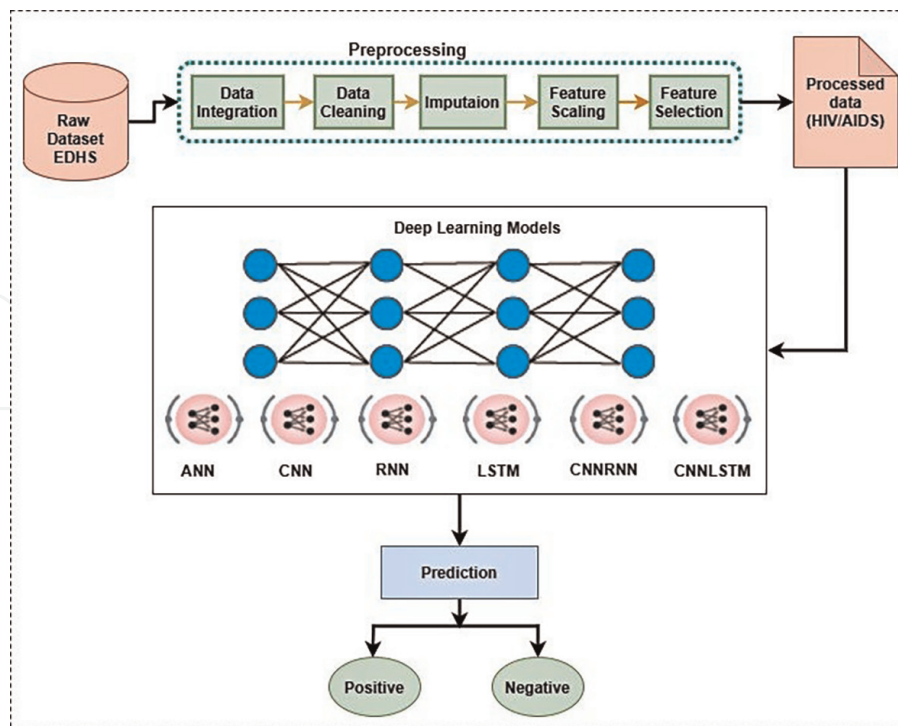
Figure 1.
The proposed deep learning approach's architecture.

3.1 Dataset

The HIV/AIDS dataset [3] has been collected from the EDHS repository from Central Statistics Agency (CSA) and DHS program https://dhsprogram.com that has been used for both training and testing purposes it is available on https://github.com/danielmesafint/Datasets. We collect this dataset as it is from the above sources and we create the HIV/AIDS dataset by considering the criteria to create a dataset from the secondary dataset, the techniques we were used to create our dataset are dataset acquisition, dataset cleaning, dataset labeling and more. The EDHS dataset has more features or attributes, it includes various demographic and health-related datasets. After the creation of our dataset, HIV/AIDS contains 83,100 instances and 33 attributes. The output level has two classes, where "0" represents Negative results (HIV-) and "1" represents Positive results (HIV+). The preprocessing section has explained how we process the EDHS dataset.

3.2 Preprocessing

As an initial step in the pre-processing stage, the original input dataset is analyzed, making the raw dataset ready for use in the prediction process [14]. In the years 2000, 2005, 2011, and 2016, four separate datasets were used to compile the dataset. The size of the dataset is reduced as a result of preprocessing. As a result, there is a scarcity of datasets, which has a negative impact on the prediction of HIV test results. As a result, the dataset integration technique is used to combine all separate dataset sets. The tight coupling method is used for integration. There are 83,100 dataset instances collected; 4,223 of these instances are incorrect, owing to user entry errors, storage or transmission corruption, or different dataset dictionary definitions of similar items in different

stores; these datasets are unreliable, inaccurate, or irrelevant. To address this issue, we use dataset cleaning techniques to identify and remove crude and incorrect instances. The dataset cleaning technique used in this process is: to remove duplicate techniques and delete all formatting techniques. After cleaning, the dataset set is uniform. There is also an issue of incompleteness with some features or variables, such as R_SeA, Had_Sex, and Con_Use, and we use the imputation technique to fill in the missing values. Some of the dataset entries in the dataset have not been completed (that is not having values present for every single variable in the dataset set). At this phase, we do two simple approaches to imputation: dropping rows with null values and dropping features with high nullity. Otherwise, the most frequent value for numerical variables and the mean for quantitative variables were used to handle missing results.

Because the nominal dataset cannot be used in a DL model, all nominal attributes, including the label class (Negative/Positive), were converted to numerical binary values with "0" and "1." The Attribute AGE is classified as 1–7. (the original value of the Age should be grouped into 1 to 7). Furthermore, depending on the form of the attribute, the unsupervised discretization filter discretized all continuous numeric attributes using different bins range accuracy. The dataset discretization technique is used to perform this transformation.

The EDHS dataset has several features, each with a unique set of numerical values, which complicates the computing procedure. As a result, a normalizing methodology is utilized to normalize dataset D^{hiv} in the range of "0" to "1", as well as to reduce numerical complexity during the HIV test status prediction computational process. Normalization may be accomplished using a variety of approaches. The well-known min-max normalizing approach is employed in the proposed system [15]. Using the following equation, this approach maps to a numeric value, D, of the initial dataset D^{hiv} into D_{norm} with an interval of [0, 1]

$$D_{norm} = \frac{D^{hiv} - D_{min}}{D_{max} - D_{min}} \: X \: [n\_max \: - n\_min \:] + n\_min \tag{1}$$

In this case, D_{norm}, D^{hiv}, D_{max}, and D_{min} represent the normalized dataset value, the original dataset value, the minimal and maximum value in the complete dataset, respectively, while n_max and n_min represent the range of the transformed dataset. We use the values n_max = 1 and n_min = 0. Using this strategy, all of the feature values fall inside the range [0, 1].

3.3 Feature selection

The EDHS-HIV/AIDS dataset were having 33 (thirty-three) variables but from these variables, we are using 20 variables as a final feature. The feature selection technique is applied to select the features. For this study, we apply the backward feature selection (BFS) technique of wrapper-based methods [16].

BFS algorithm aims to reduce the dimensionality of the initial feature subspace from N to *K-features* with a minimum reduction in the model performance to improve upon computational efficiency and reduce generalization error. The primary idea is to sequentially remove features from the given features list consisting of N features to reach the list of *K-features*. At each stage of removal, the feature that causes the least performance loss gets removed.

We use the hit and trial method for different values of *K-features* and evaluate all subsets of features using their obtaining accuracy and making the final decision. Based

on this, we select the 20 best features and the selected features are presented in **Table 1**. Moreover, the selected feature helps to reduce the over-fitting of the DL models, making the training time fast, reducing the complexity, and easier to interpret our models, and then it helps to make a better prediction power.

After preprocessing we use a total of 78,877 (83,100–4223) instances and 20 (from 33) features. **Table 1** shows the statistical descriptions of the selected features.

3.4 Deep learning models

The study's goal is to create an HIV test status prediction model by employing six deep learning models that have not been used before in HIV test result prediction. Recently, different deep learning techniques and their combinations are widely used for demographic and health dataset prediction or classification based on some obtained parameters.

In this work, we create and test prediction models for HIV status based on demographic and health survey datasets. To assess the study, we trained four DL models, including Artificial Neural Network (ANN), Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Long Short Term Memory (LSTM), and two hybrid DL Models such as CNNRNN and CNNLSTM.

Features	MIN	MEAN	MAX	Descriptions
Sex	1.000000	1.502419	2.000000	Gender of the person
Age	1.000000	3.259434	7.000000	Age of the person
Reg	1.000000	5.177661	15.000000	Region where is lived
M_Sta	0.000000	0.885908	5.000000	Marital status of a person
W_Ind	1.000000	3.196857	5.000000	Standard of Living
H_Sex	1.000000	1.222080	2.000000	Did you have Sex?
R_SeA	0.000000	0.607065	1.000000	Recent sexual activity
N_S_Part	1.000000	1.121004	3.000000	How many sex partners do you have?
C_Use	0.000000	0.346349	1.000000	Can you use a condom?
R_Use_Con	0.000000	0.762203	1.000000	Refuse to use a condom?
R_Nhave_Sex	0.000000	0.735561	1.000000	Refuse not to have sex?
HIV_Mosq	0.000000	0.401831	1.000000	Did get HIV by Mosquito?
H_STI	0.000000	0.285557	1.000000	Did you hear HIV transmission?
H_O_STI	0.000000	0.317876	1.000000	Did you hear other means of transmission?
H_AIDS	0.000000	0.966567	1.000000	Did you hear about AIDS?
E_T_HIV	0.000000	0.399634	1.000000	Did you test HIV before?
P_T_HIV	0.000000	0.384085	1.000000	Where do you test?
S_Test	0.000000	0.244282	1.000000	Sample test?
T_in_LAB	0.000000	0.914153	1.000000	Laboratory test
F_T_Resu	0.000000	0.203449	1.000000	Final results

Table 1.
The statistical descriptions of the selected features using backward feature selection methods.

ANN [17] is commonly used for prediction and modeling tasks. Because of its self-learning and self-adapting abilities, ANN is an interesting choice for estimating underlying dataset relationships. It is made up of different neurons, input, output, hidden layers, and activation functions. CNN [18] is a neural network type that is often utilized in image categorization research. It has layers like pooling, convolutional, classification, and fully-connected layer. CNN, contrasting ML, acquires characteristics on its own. The dimension of the inputs is lowered in the pooling layer. RNN [19] is a type of feed-forward NN that includes internal memory. RNN employs the same procedure for each input; however, the result of the input dataset is reliant on the previous result. RNN processes inputs using its internal memory. LSTM [19] is a variant of the RNN. It is simpler to recall the previous dataset in the LSTM. The LSTM networks address the RNN vanishing gradient problem. CNNRNN [20] is a hybrid model that uses a different convolutional layer and a single recurrent layer to process the input sequence of characters. CNNLSTM [21] is a hybrid of CNN and LSTM layers that provides the benefits of both models.

3.5 Performance evaluation metrics

Before building a prediction model, all models must be assessed using several evaluation parameters [22]. We've so far used accuracy scores to evaluate our prediction models. But sometimes accuracy score isn't all enough to evaluate a model properly as the accuracy score doesn't tell exactly which class (positive or negative) is being wrongly predicted by our models in case of a low accuracy score. To clarify this, we perform precision score; recall score, f1 score, AUC, and log-loss for both models. And then we compare our models using these calculated metrics to see exactly where one model excels over the other. We utilized 10-fold CV and an 80:20 train-test split technique to validate the utilized dataset.

$$\text{Acc} = \frac{TP + TN}{TP + TN + FP + FN} \tag{2}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{3}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{4}$$

As shown in Eqs. (2)-(4), where true positive (TP) is the number of HIV-positive persons who are actually positive. The number of predicted negative persons that are actually negative is represented by the true negative (TP). The amount of people who are labeled as positive but are actually negative is known as false positive (FP). The number of labeled negative persons who are actually positive is defined as false negative (FN). These metrics are frequently computed to measure the predictive quality of models.

$$\text{F1} - \text{score} = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{5}$$

The F1 score may be calculated by dividing the product of recall and precision by the total of recall and precision, as shown in Eq. (5).

$$\text{AUC} = \int_{-\infty}^{\infty} y(t)dx(t) \tag{6}$$

$$\text{Log loss} = (x+a)^n = \sum_{i=0}^{n} y_i \bullet (\log(P_i) + (1-y) \bullet \log(1-P_i)) \tag{7}$$

As Eq. (7), where n is the samples count, yi is the label of the actual class, and pi is the probability of ith sample fits one class. The model performance is measured using log-loss, which computes the prediction as a probability value between "0" and "1". A better predictor must have a lower error value of log-loss, for the goal of lowering it to "0" in the case of a perfect predictor.

4. Experiment setups and result discussion

This part presents the experimental setting and experimental findings and analysis.

4.1 Experimental setting

We carried out our model experiments on Microsoft Windows 10 with an Intel® Core™ i7- 9700 CPU running at 3.00 GHz, 8 processors, 16 GB RAM, and a 1 TB hard disc. The Python language version 3.6 tool with Keras [23] and Tenser-flow was utilized.

To evaluate a model's performance, we need some dataset (input) for which we know the ground truth (label). For this problem, we don't know the ground truth for the test set but we do know for the train set. So the idea is to train and evaluate the model performance on HIV/AIDS dataset. One thing we do is to split the train set into two groups, in the case we use the 80:20 ratio, the ratio is done randomly. That means we would train the model on 80% of the training dataset and we reserve the rest 20% for evaluating the model since we know the ground truth for this 20% dataset. Then we compare our model prediction with this ground truth (for 20% dataset). That's we observe how our model would perform on the unseen dataset. This is the first model evaluation technique. This process is used by the sklearn library in a train-test split method [24].

The parameters setting: For ANN we have 3 hidden dense layers with 32, 16, and 8 perceptrons and the last layer is activation functions with sigmoid. For CNN we have 2 hidden CNN layers with 512 and 256 perceptrons with MaxPooling1D function followed by 2 fully connected layers with 2048 and 1024 perceptrons. And the last layer is activation functions with the sigmoid. For RNN the input layer is Simple RNN with 512 perceptrons followed by 2 fully connected layers with 2048 and 1024 perceptrons. And the last layer is activation functions with the sigmoid. For LSTM the input layer is LSTM with 512 perceptrons followed by 2 fully connected layers with 2048 and 1024 perceptrons. And the last layer is activation functions with the sigmoid. For CNNLSTM the input layer is Conv1D with 512 perceptrons followed by MaxPooling1D layer, and the output of them is connected to the LSTM layers with 512 perceptrons. And the last hidden layers are dense layers with 2048 and 1024 perceptrons. For CNNRNN the input layer is Conv1D with 512 perceptrons followed by the MaxPooling1D layer, and the output of them is connected to the RNN layers with 512 perceptrons. And the last hidden layers are dense layers with 2048 and 1024 perceptrons. Batch Normalization and Dropout layers have been added to all the

models to improve the accuracy and help to avoid overfitting. For all models the Learning rate is 0.001, the Loss function is Binary Cross entropy, the Decay is 0.0001, and the optimizer is ADAM [25].

4.2 Experimental result analysis

This subsection presents a detailed analysis of the experimental findings achieved using the proposed approach on HIV/AIDS datasets with standard performance metrics.

As a predictor, six DL models were constructed and used. Predictions were then made, and the performance was assessed. The first experiment is conducted with a train-test split.

For performance evaluation, in terms of goodness-of-fit, the HIV test result prediction model performances are compared. The model compared in this proposed method is; the RNN model, achieving an accuracy of 0.870, the precision of 0.871, recall of 0.876, f1-score of 0.876, and AUC of 0.94. As shown in **Table 2**, all DL models' accuracy results were at least 0.834 or above. With 0.870, the RNN model had the best evaluation performance. RNN was implemented considering several parameters such as dropout, batch-size, epochs, optimizers, etc. The performance of the RNN was based on those parameters. Thus, the performance of RNN is slightly better than the other DL models. The CNNLSTM hybrid model was shown to be the second-best model with 86.2%.

Performance measure metrics values were found to be more than 83.0%. Precision is the proportion of accurately predicted positive findings to the total number of expected positive findings. A perfect precision in information retrieval experiments should be 1. The greatest precision score in this study was obtained using RNN, which was 0.871. A ratio of accurately predicted positive findings to all results is defined as recall. A recall score, like accuracy, must be one for the categorization process to be perfect. With 0.876, the best recall value was attained using the RNN model. F1-score calculated as the weighted average of accuracy and recall scores. This criterion considers both FP and FN. A high F1-score indicates that the predictor has few FP and few FN. In this scenario, the predictor identifies serious threats while avoiding false alarms. When the value of an F1-score is 1, it is deemed perfect. The best F1-score got with RNN was 0.876, as with any other assessment criterion. In classification analysis, The AUC is used to determine the best algorithms used to predict target classes. In general, a score value of AUC 0.5 indicates that no variance, a score between 0.6 and 0.8 is held as allowable, a score of 0.8–0.9 is regarded as excellent, and a value of

Model	Accuracy in training	Accuracy of testing	Precision	Recall	F1-score	AUC	Log-loss
ANN	0.881	0.8629	0.858	0.858	0.858	0.94	0.3137
CNN	0.879	0.856	0.856	0.856	0.856	0.94	0.3201
RNN	0.910	0.870	0.871	0.876	0.876	0.94	0.2909
LSTM	0.870	0.844	0.844	0.844	0.844	0.92	0.3587
CNNRNN	0.881	0.862	0.862	0.862	0.862	0.93	0.3130
CNNLSTM	0.863	0.834	0.835	0.835	0.835	0.93	0.3707

Table 2.
*The evaluation outcomes of all DL models using the train-test split method.*

greater than 0.9 is regarded as exceptional [26]. The AUC values of all DL models were outstanding since all of the outcomes were more than 0.9. All DL models may be used to predict HIV test results based on their AUC values.

True positive rates are critical in health investigations since recall indicates the percentage of actual positives identified [27]. A recall is a significant assessment criterion in this study since it is computed by dividing the number of properly-recognized HIV-positive samples by the total number of HIV test results. Besides, the AUC score plays an important role in health research since it has a relevant interpretation for health prediction [28]. Accuracy is a study criterion that indicates how near the sample parameters are to population characteristics. We can demonstrate that the study is generalizable, dependable, and valid by testing the correctness of the models [29]. As a result, just these three assessment indicators were examined in this study. The remaining ones were computed to compare the findings to earlier studies. The AUC values using the train test split strategy are shown in **Figures 2**–**7**.

In addition to the metrics listed above, we calculated prediction accuracy to assess the efficacy of the proposed approach. **Figures 8**–**13** depict the prediction accuracy of the proposed technique on the HIV/AIDS dataset in terms of each DL model. Because of
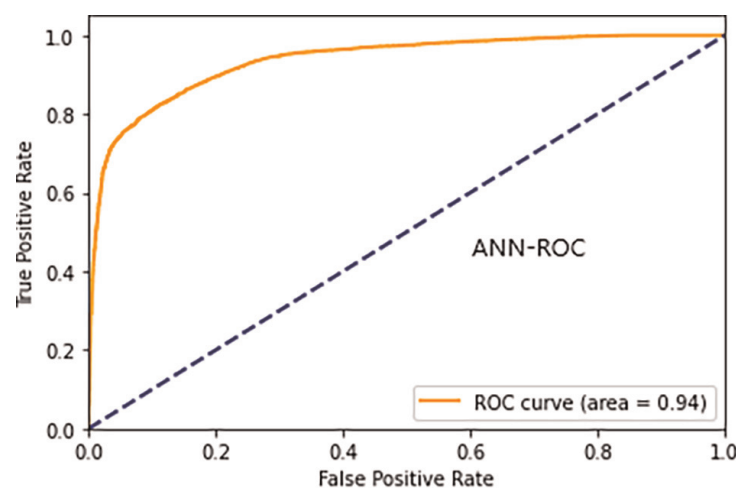
Figure 2.
ANN models AUC using the train-test split strategy.
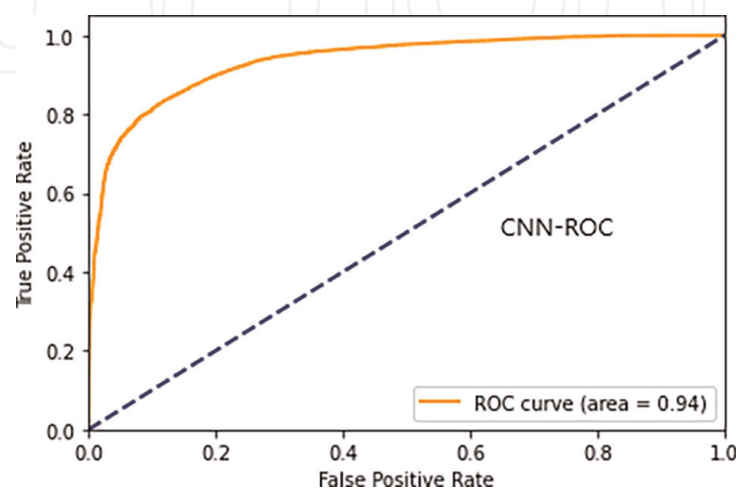
Figure 3.
CNN models AUC using the train-test split strategy.
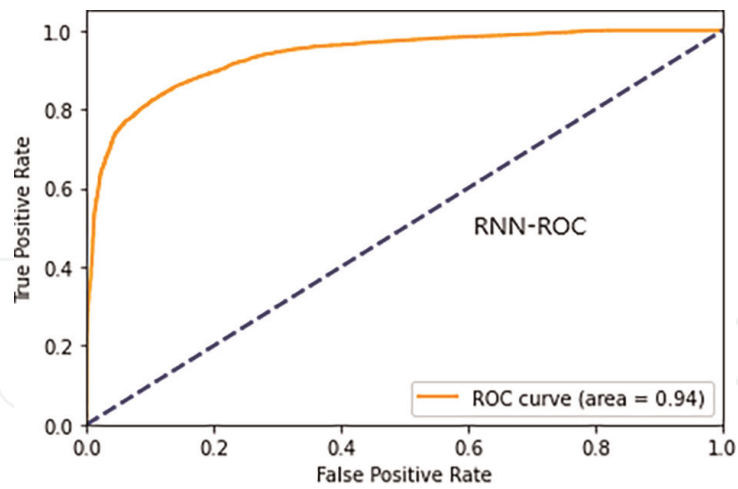
Figure 4.
RNN models AUC using the train-test split strategy.
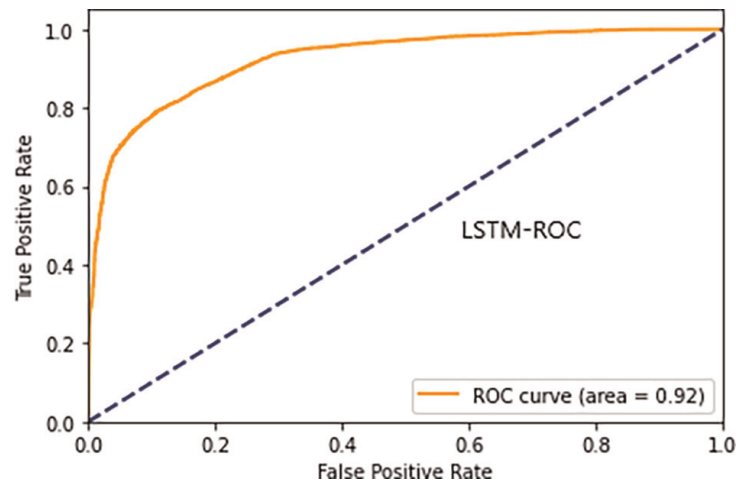
Figure 5.
LSTM models AUC using the train-test split Strategy.
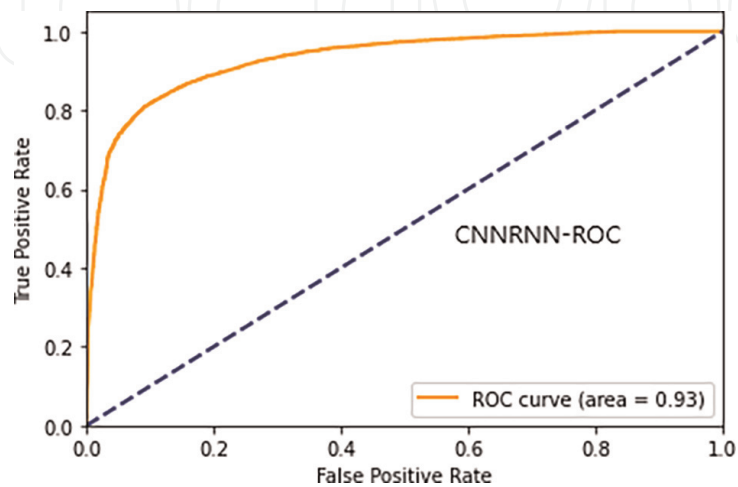
Figure 6.
*CNNRNN models AUC using the train-test split strategy.*

11
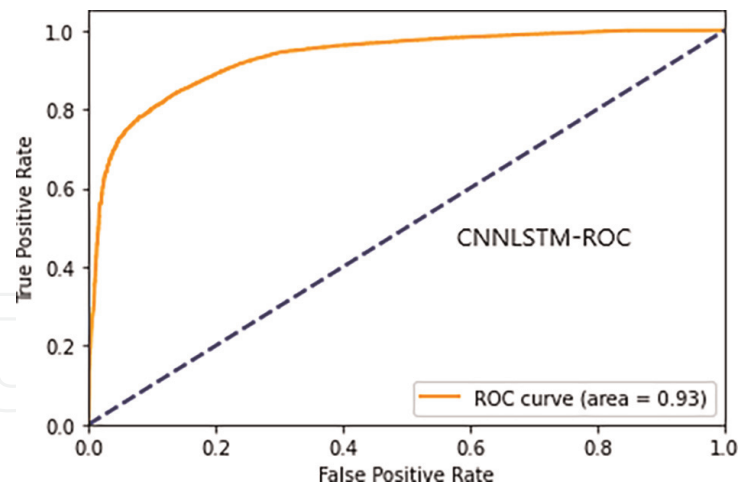
Figure 7.
CNNLSTM models AUC using the train-test split strategy.

the flawless prediction accuracy of the HIV/AIDS dataset, a substantial difference is not there among lines related to the training and test dataset, as shown in **Figures 8–13**.

We also used the log-loss error function to evaluate our work. As shown in **Figures 14–19** the training sample loss is near to 0, while the stated loss with the test sample is 0.3707 (refer to **Table 2**) implying that more research on this specific dataset is required to reduce the error. As demonstrated in Loss Figures, the proposed technique outperforms all DL models by scoring the least number of errors in the test instances, with the ANN, CNN, RNN, LSTM, CNNLSTM, and CNNRNN scoring 0.3137, 0.3201, 0.2929, 0.3587, 0.3130, and 0.3707, correspondingly (refer **Table 2**). It is noticed that a distance between the training and the test lines indicates whether or not the model is over-fitting.

The second experimental result for this work is a 10-fold CV. **Table 3** demonstrates the assessment results of all DL models using a 10-fold CV technique.

Concerning the predictive performances, we discovered that the best comprehensive recognized models on AUC score for predicting HIV test status were 89.72 by ANN. The main reason behind ANN outperforming better results is its activation function unlike CNN, RNN, and LSTM. Moreover, ANN works better for numerical datasets unlike CNN, RNN, and LSTM which work on image data and time-series data
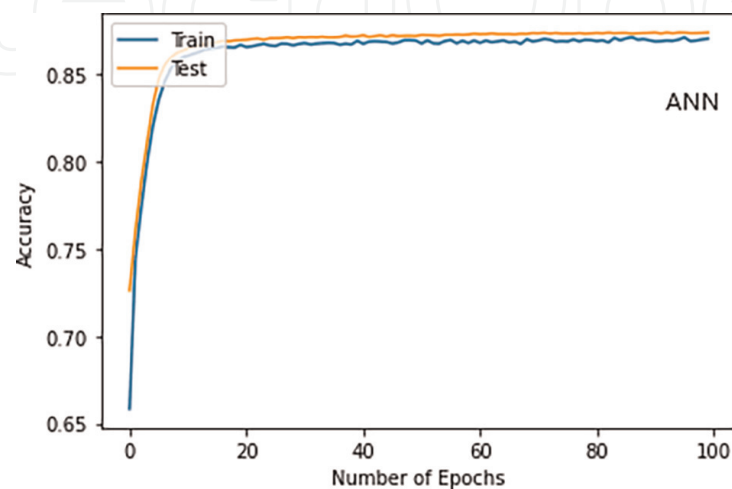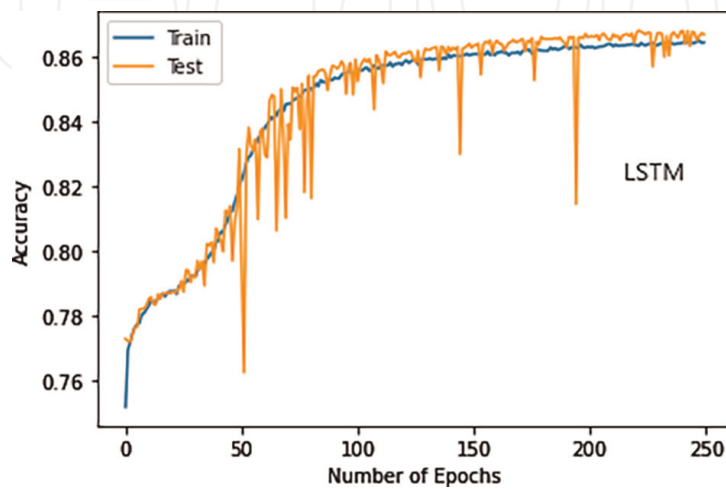

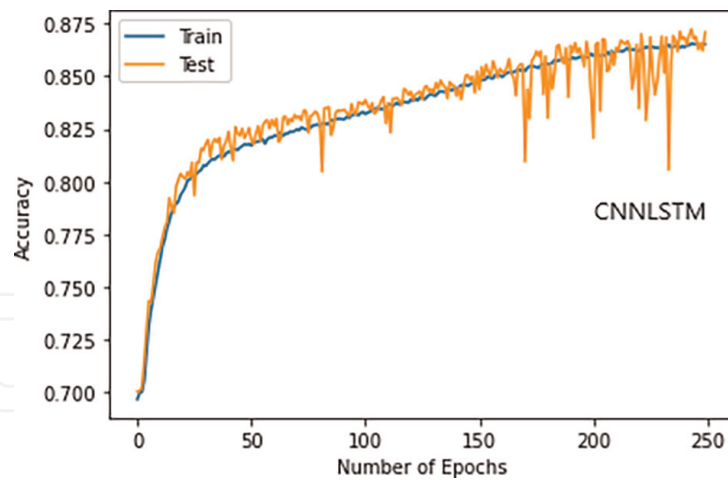**Figure 8.**
The prediction accuracy on the model ANN.
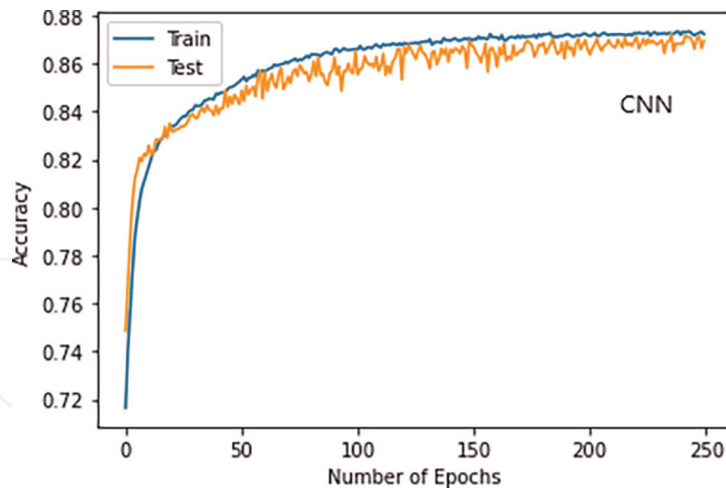
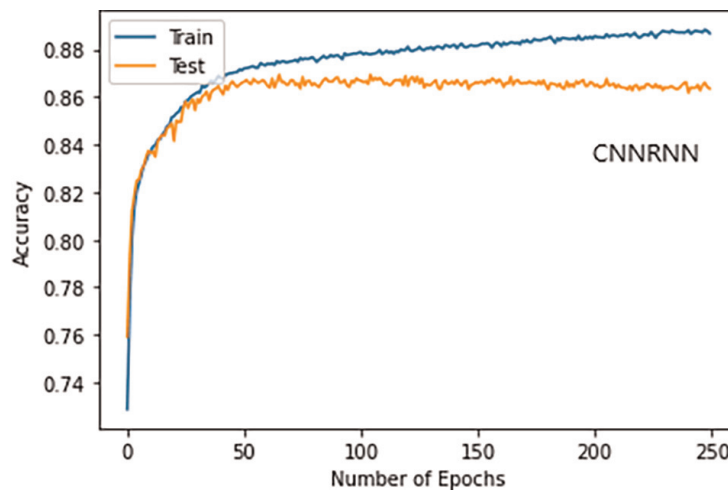**Figure 9.**
The prediction accuracy on the model CNN.
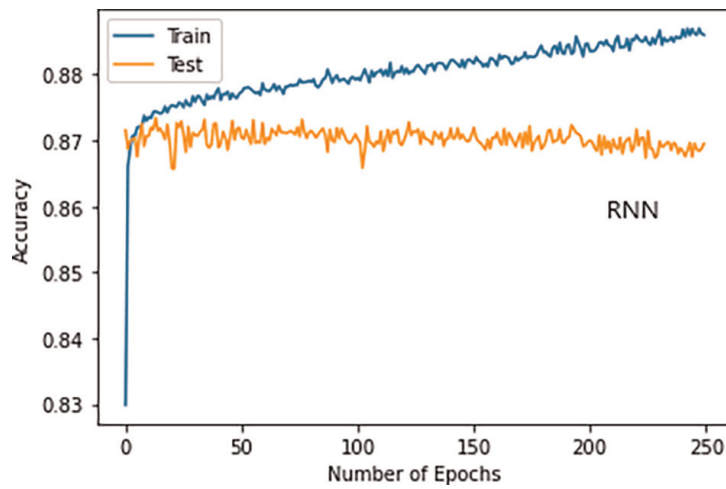
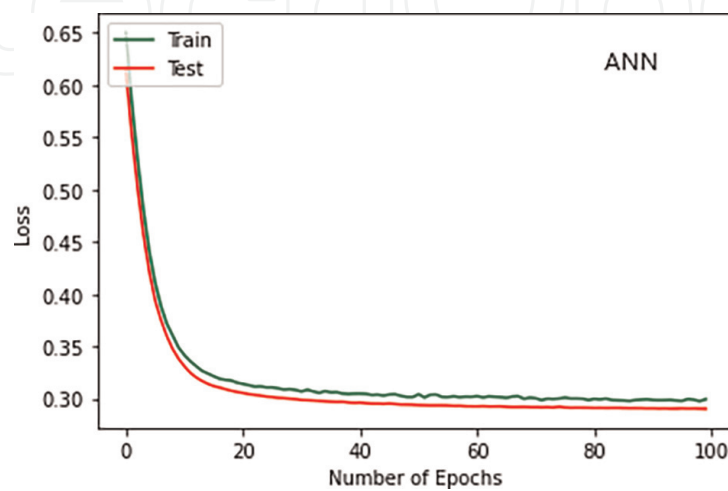**Figure 10.**
The prediction accuracy on the model RNN.

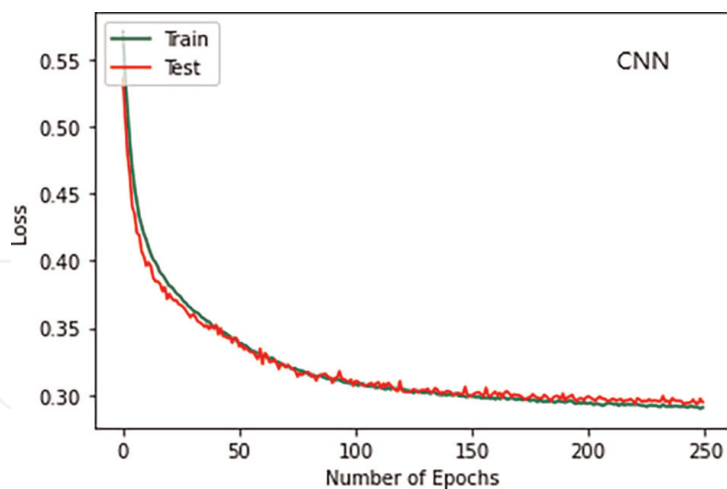**Figure 11.**
The prediction accuracy on the model LSTM.

Figure 12.
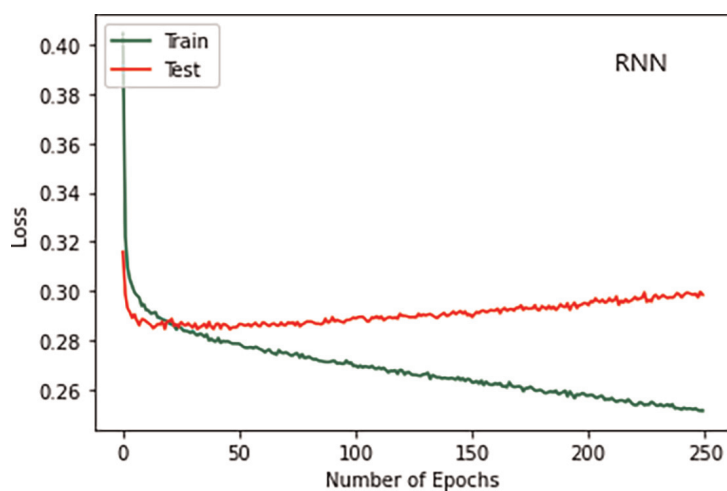*The prediction accuracy on the model CNNLSTM.*

Figure 13.
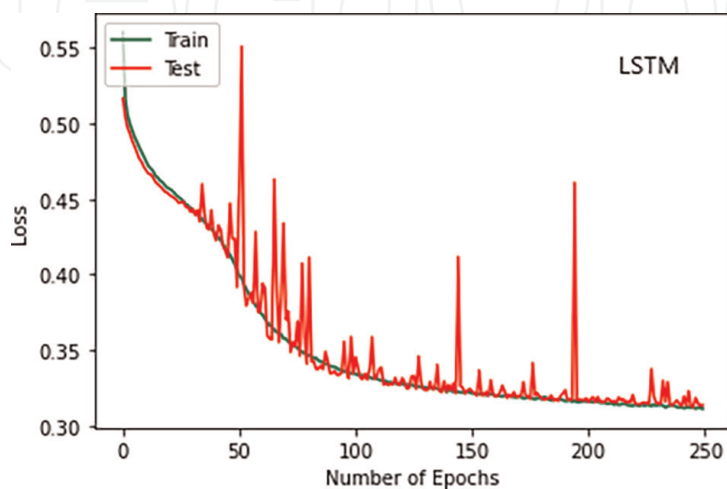*The prediction accuracy on the model CNNRNN.*

Figure 14.
Prediction Loss on the model ANN.

Figure 15.
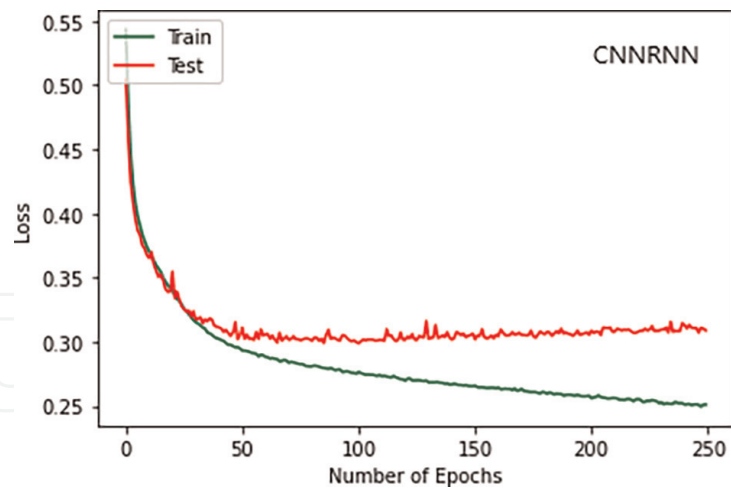*Prediction loss on the model CNN.*

Figure 16.
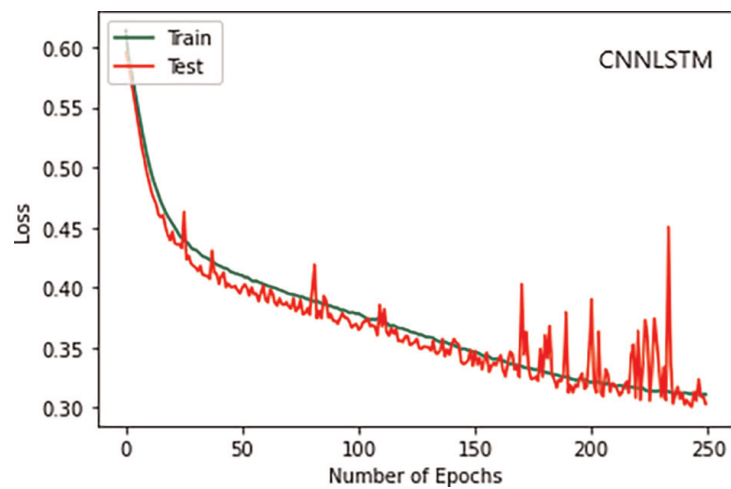*Prediction loss on the model RNN.*

Figure 17.
Prediction loss on the model LSTM.

Figure 18.
Prediction loss on the model CNNRNN.

Figure 19.
Prediction loss on the model CNNLSTM.

Model	Accuracy	Precision	Recall	F1-score	AUC
ANN	0.855	0.844	0.857	0.851	89.72
CNN	0.834	0.833	0.845	0.841	87.52
RNN	0.847	0.821	0.822	0.834	87.62
LSTM	0.821	0.830	0.836	0.831	86.46
CNNRNN	0.801	0.800	0.822	0.826	83.78
CNNLSTM	0.792	0.800	0.821	0.801	83.77

Table 3.
The outcomes of all DL models were evaluated using a 10-fold cross-validation methodology.

respectively. It was discovered that predicting HIV test status from the EDHS dataset was considered a difficult activity. Nonetheless, the best HIV test status prediction outcomes using ANN obtained reasonable accuracy of 85.5%, precision of 84.4%, recall of 85.7%, and f1-score of 85.1%.

References	Dataset	Methods	Best Results	Accuracy	F1-score	AUC
McSharry et al. [8]	Population-based HIV Impact Assessment	Machine Learning	XGBoost	—	0.789	—
Orel et al. [9]	DHS dataset	Machine Learning	SVM	0.80	—	—
Ahlström et al. [12]	Danish National Hospital Registry	Machine Learning	LR	—	—	88.4
Lu et al. [13]	UCI ML repository	Deep Learning		—	0.927	—
Steiner et al. [22]	HIV Drug Resistance Dataset	Deep Learning	MLP	0.826	0.732	0.90
Betechuoh et al. [30]	Antenatal Survey	Deep Learning	ANN	0.840	—	0.86
Proposed method	EDHS –HIV/AIDS	Deep learning	RNN	0.87	0.87	0.94

Table 4.
The proposed model comparison with some of the most recent related research works.

5. Comparison

This section of the study compares the proposed method to certain selected recent research in terms of performance measures. **Table 4** compares the proposed method's assessment metrics to those of six other recent research works. The hyphen (-) in the table's specific cells indicates that the researchers did not consider metrics in their study. As shown in **Table 4**, the best results were obtained with various models. Nonetheless, we have not employed ML in our research. We created six DL models and achieved higher accuracy, f1-scores, and AUC when compared to earlier similar efforts. In the considered HIV/AIDS dataset, the suggested technique obtains improved prediction performance, with 0.87 in total accuracy and f1-score and 0.94 in AUC score.

6. Conclusion

In this work, deep learning models based on the EDHS dataset were used to predict HIV test results. Six deep learning models were used to analyze HIV/AIDS dataset. The dataset was normalized in the first stage of the study then utilized as an input for the DL models. Following that, prediction is performed, and the models' results were evaluated using precision, recall, accuracy, AUC, and F1-scores. We used 10 fold CV and train test split techniques to assess the models. In a 10-fold CV technique, the ANN deep learning model produced the most meaningful results, with an accuracy of 85.5%, a recall of 85.7%, and an AUC score of 87.72%. Despite its popularity, this validation did not produce the best validation results. In the train-test split technique, the greatest accuracy, precision, recall, and AUC values were obtained with the RNN model, which was 87%, 87%, 87%, and 94%, respectively. The accuracy of all DL models produced in the study was greater than 83%. Precision and recall values can be inferred in the same way.

Finally, we discovered evidence that DL models may be used to predict HIV test status using demographic and health survey datasets. Our findings on the role of DHS

in predicting HIV test status for people improve our knowledge of the consequences of HIV epidemics. Based on the findings of our study, we believe that the health domain should investigate the use of DL models that analyze individual HIV test status to enhance and re-evaluate health policies and intervention mechanisms.

Acknowledgements

Funding

Author contributions

This is a collaborative work with both authors that contribute throughout.

Conflict of interest

The authors declare that they have no conflict of interest.

Ethical standard

This article does not contain any studies with human participants or animals performed by any of the authors.

Data availability

The authors declare that all data supporting the findings of this study are available on https://github.com/danielmesafint/Datasets

Author details

Daniel Mesafint Belete* and Manjaiah D. Huchaiah
Department of Computer Science, Mangalore University, India

*Address all correspondence to: danielmesafint1985@mail.com;
drmdhmu@gmail.com

IntechOpen

References

[1] WHO. HIV/AIDS fact sheet. 2017. Available from: http://www.who.int/features/factfiles/hiv/en/

[2] Huerga H et al. Who needs to be targeted for HIV testing and treatment in KwaZulu-Natal? Results from a population-based survey. Journal of Acquired Immune Deficiency Syndrome. 2016;**73**(4):411-418. DOI: 10.1097/QAI.0000000000001081

[3] CSA, Demographic and Health Survey. 2018. [Online]. Available from: http://www.csa.gov.et/survey-report/category/2-demographic-and-health-survey [Accessed: October 28, 2018]

[4] Nassif AB, Shahin I, Attili I, Azzeh M, Shaalan K. Speech recognition using deep neural networks: A systematic review. IEEE Access. 2019;**7**:19143-19165

[5] Ramabhadran B, Khudanpur S, Arisoy E. Will We Ever Really Replace the N-gram Model? On the Future of Language Modeling for HLT. In: Proceedings of the NAACL-HLT, Montreal. Canada: Omni Press Inc.; 2012. pp. 1-10

[6] Aljunid MF, Huchaiah MD. Multi-model deep learning approach for collaborative filtering recommendation system. CAAI Transactions on Intelligence Technology. 2020;**5**(4):268-275

[7] Ciregan D, Meier U, Schmidhuber J. Multi-column deep neural networks for image classification. In: IEEE Conference on Computer Vision and Pattern Recognition. Providence, RI, USA: IEEE; 2012. pp. 3642-3649

[8] McSharry PE, Mutai C, Ngaruye I, Musabanganji E. Use of machine learning techniques to identify HIV predictors for screening in sub-Saharan Africa. BMC Medical Research Methodology. 2021;**1**:1-11

[9] 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. BMJ. 2020:1-29

[10] Lu X, Wang L, Jiang Z. The application of deep learning in the prediction of HIV-1 protease cleavage site. In: 5th International Conference on Systems and Informatics (ICSAI). Nanjing, China: IEEE; 2018. pp. 1299-1304

[11] Wang G, Wei W, Jiang J, Ning C, Chen H, Huang J, et al. Application of a long short-term memory neural network: A burgeoning method of deep learning in forecasting HIV incidence in Guangxi, China. Epidemiology & Infection. 2019;**147**(194):1-7

[12] Ahlstrom MG, Ronit A, Omland LH, Vedel S, Obel N. Algorithmic prediction of HIV status using nation-wide electronic registry dataset. E Clinical Medicine. 2019;**17**:100203

[13] Steiner MC, Gibson KM, Crandall KA. Drug resistance prediction using deep learning techniques on HIV-1 sequence dataset. Viruses. 2020;**12**(5):560

[14] Garcia S, Luengo J, Herrera F. Dataset Preprocessing in Dataset Mining. Intelligent Systems Reference Library book series. Vol. 72. Singapore: Springer; 2015

[15] Jain YK, Bhandare SK. Min max normalization based dataset perturbation method for privacy protection. International Journal of Computer & Communication Technology. 2011;**2**(8):45-50

[16] Manjaiah D, Belete DM. Wrapper based feature selection techniques on EDHS-HIV/AIDS dataset. European Journal of Molecular& Clinical Medicine. 2020;**7**(8):2642-2657

[17] Han J, Kamber M, Pei J. Dataset mining concepts and techniques third edition. The Morgan Kaufmann Series in Dataset Management Systems. 2011; **5**(4):83-124

[18] Wu J. Introduction to Convolutional Neural Networks. Vol. 5, No. 23. China: National Key Lab for Novel Software Technology, Nanjing University; 2017. pp. 1-30

[19] Sherstinsky A. Fundamentals of recurrent neural network (rnn) andlong short-term memory (lstm) network. Physica D: Nonlinear Phenomena. 2020; **404**:132306

[20] Xiao Y, Cho K. Efficient character-level document classification by combining convolution and recurrent layers. Computer Science - Computation and Language. 2016;**65**:1-10

[21] Rahman M, Islam D, Mukti RJ, Saha I. A deep learning approach based on convolutional LSTM for detecting diabetes. Computational Biology and Chemistry. 2020;**88**:107329

[22] Xie Y, Zhu C, Zhou W, Li Z, Liu X, Tu M. Evaluation of machine learning methods for formation lithology identification: A comparison of tuning processes and model performances. Journal of Petroleum Science and Engineering. 2018;**160**:182-193

[23] Chollet F. Keras. 2018. Available from: https://keras.io/ [Accessed: June 10, 2021]

[24] Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine learning in python. Journal of machine Learning Research. 2011;**12**: 2825-2830

[25] Kingma DP, Ba J. Adam: A method for stochastic optimization. In: 3rd Int. Conf. for Learning Representations. Vol. 1. 2014. pp. 1-15

[26] Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. Journal of Thoracic Oncology. 2010;**5**(9):1315-1316

[27] Avati A, Jung K, Harman S, Downing L, Ng A, Shah NH. Im-proving palliative care with deep learning. BMC Medical Informatics and Decision Making. 2018;**18**(4):55-64

[28] Kamarudin AN, Cox T, Kolamunnage-Dona R. Time-dependentroc curve analysis in medical research: Current methods and applications. BMC Medical Research Methodology. 2017;**17**(1):1-19

[29] Pierce R. Evaluating information: Validity, reliability, accuracy, triangulation. In: Research Methods in Politics: A Practical Guide. Edmonton, AB, Canada: Sage Publications; 2008

[30] Betechuoh BL, Marwala T, Tettey T. Autoencoder networks for HIV classification. Current Science. 2006; **91**(11):1467-1473