BAREKENG: Journal of Mathematics and Its Applications
June 2026 Volume 20 Issue 2 Page 1473–1484
P-ISSN: 1978-7227 E-ISSN: 2615-3017
https://doi.org/10.30598/barekengvol20iss2pp1473-1484

IMPLEMENTATION OF KALMAN FILTER, RECURRENT NEURAL NETWORK, AND DECISION TREE METHOD TO FORECAST HIV CASES IN EAST JAVA

Nurwijayanti[1], Firman Yudianto[2], Panca Radono[3], Rizky Amalia Sinulingga[4*], Mochammad Romli Arief[5], Rahayu Budi Utami[6], Hamzah Arof[7]

[1] Public Health Department, Faculty of FAKAR, Universitas STRADA Indonesia
[3]Hospital Administration Department, Faculty of FAKAR, Universitas STRADA Indonesia
Jln. Manila No. 37, Kediri, 64132, Indonesia

[2] Department of Information Systems, Faculty of Economics, Business and Digital Technology, Universitas Nahdlatul Ulama Surabaya
[5] Undergraduate Program of Information System, Universitas Nahdlatul Ulama Surabaya
Jln. Raya Jemursari 51-57 Surabaya, 60237, Indonesia

[4] Department of Business, Faculty of Vocational, Universitas Airlangga
Jln. Dharmawangsa Dalam Selatan 28-30, Surabaya, 60286, Indonesia

[6]Nursing Science Study Program, Sekolah Tinggi Ilmu Kesehatan Satria Bhakti
Jln. Brantas No 3B, Babadan, Werungotok, Nganjuk, 64412, Indonesia

[7]Department of Electrical Engineering, University of Malaya
Kuala Lumpur, 50603, Malaysia

*Corresponding author's e-mail: ** **rizkyamalia@vokasi.unair.ac.id**

Article Info

Article History:

Received: 7th June 2025
Revised: 12th August 2025
Accepted: 24th September 2025
Available online:26th January 2026

Keywords:

Forecasting;
HIV;
Machine learning.

ABSTRACT

HIV (Human Immunodeficiency Virus) is a virus that infects cells in the body and weakens the human immune system, making it more susceptible to various diseases. Meanwhile, the symptoms of the disease arising from HIV itself are referred to as AIDS (Acquired Immune Deficiency Syndrome). Approximately 50% of people with AIDS in Indonesia are adolescents. Until now, HIV/AIDS has ranked second in East Java province. HIV/AIDS is classified as a dangerous disease because of the risk of death. Unfortunately, there is no treatment method or vaccine that could prevent this disease. This monitoring program to prevent the development of dangerous health cases such as HIV/AIDS is very helpful for local governments. Along with the development of information technology, the emergence rate of new HIV/AIDS cases can now be forecasted using machine learning as a monitoring tool to support. This machine learning-based monitoring program works with past data for statistical analysis. In this study, the methods used are Kalman Filter, Recurrent Neural Network, and Decision Tree. The Kalman Filter is a type of filter method that is used to predict the state of a dynamic, stochastic, linear, discrete system. A Recurrent Neural Network (RNN) is a development of a Neural Network. RNN deals with input sequence/time-series data by individual sector at each step and preserves the information it has captured at previous time steps in a hidden state. A Decision Tree is one of the classic tree-based prediction methods. The best error value (RMSE) achieved by each method is 0.0885 for the Kalman Filter, then for the Recurrent Neural Network method achieved 0.198, and the Decision Tree method successfully achieved 0.0287.

How to cite this article:

Nurwijayanti, F. Yudianto, P. Radono, R. A. Sinulingga, M. R. Arief, R. B. Utami, and H. Arof, "IMPLEMENTATION OF KALMAN FILTER, RECURRENT NEURAL NETWORK, AND DECISION TREE METHOD TO FORECAST HIV CASES IN EAST JAVA", *BAREKENG: J. Math. & App.*, vol. 20, no. 2, pp. 1473-1484, Jun, 2026.

Journal homepage: *https://ojs3.unpatti.ac.id/index.php/barekeng/*
Journal e-mail: *barekeng.math@yahoo.com; barekeng.journal@mail.unpatti.ac.id*

Research Article · Open Access

1. INTRODUCTION

HIV continues to spread globally. HIV (Human Immunodeficiency Virus) is a type of virus that infects white blood cells and also weakens the human immune system [1]. HIV/AIDS has been added to the list of deadly infectious diseases, but until today, health researchers and practitioners still have not found a medicine or vaccine [2]. The World Health Organization (WHO) predicts that 36.7 million people across the globe are living with HIV/AIDS at the end of 2016 [3]. In Asia and the Pacific, 30% of new HIV infections occurred among MSM in 2018. Eastern Europe and Central Asia have been the regions with the highest incidence rates of new HIV cases over the previous decade, with 100,000 new HIV infections in 2019 [4]. In 2020, the new number of HIV infections reached 1.5 million globally, with a total of 79.3 million active cases [5]. High HIV prevalence among MSM was found in four countries, namely Thailand, Indonesia, Malaysia, and Vietnam [6].

Indonesia is one of the countries in the world with a significant improvement in HIV/AIDS cases. HIV/AIDS becomes the top program priority for prevention [7]. DKI Jakarta becomes the province with the highest number of HIV (Human Immunodeficiency Virus) with a number of 55,099 cases, followed by East Java with 43,399, West Java with 31,293, Papua with 30,699, and the last is Central Java with 24.757 [8]. Since 2010, several provinces in Indonesia have been working with the Social Ministry of Indonesia to eliminate prostitution to reduce the number of HIV infections [9]. HIV/AIDS is contracted through contact or ingestion of infected bodily fluids. Bodily fluids that can mediate HIV transmission include [10]. Therefore, serious attention is needed as a response to tackle them in order not to impact a healthy society. One of the most important challenges for effective control of HIV and AIDS transmission is stigma and discrimination [11].

Forecasting epidemics like HIV and modelling their outcomes greatly implicates health systems and policymakers [12], in this case, especially for the East Java province government, and also for specific health organizations to design preventive and strategic solutions. Currently, the use of machine learning technology to forecast HIV cases has been widely used. Machine learning entails the utilization of computational and statistical algorithms to determine hidden associations of data that might increase prediction through relaxation of the modelling postulates advanced by standard approaches. Among the latest advances in prediction approaches and identification methods in HIV statistical data, machine learning offers greater capability in processing large amounts of data [13].

In this study, the authors applied the Kalman Filter, Recurrent Neural Network (RNN), and Decision Tree methods to compare the accuracy level of each method. The Kalman Filter is a popular filter method that is used to study multivariable systems, highly volatile data, and time-varying systems [14] from a dynamic stochastic linear discrete system. The Kalman Filter is an algorithm used to update the prediction of a variable that is not directly measured from observation data [15]. RNNs are the most widely used NN architecture for sequence prediction issues [16]. RNNs are a very promising method due to the internal memory that can remember the important features of the input sequential data, which allows them to accurately predict the future [17]. RNN belongs to the domain of Deep Learning (DL). The recurrent connection to the same neurons in the prior time step, along with their internal state (memory), makes RNNs well-suited for modeling temporal behavior [18]. The last algorithm used in this study is the Decision Tree. Decision Tree belongs to the tree-based method that handles multi-output cases with little data preparation [19]. Decision Tree (DT) method, belonging to the supervised learning class algorithm, is mostly preferred for completing classification cases, but either way, it can be used in classifying as well as in regressing cases [20]. In the Decision Tree method, very large data samples will be represented as smaller orders [21].

2. RESEARCH METHODS

2.1 Literature Review

The findings and their resources for the literature review are shown in **Table 1** below.

Table 1. Literature Review

Source	Title	Key Findings
Herlambang *et al.*	Comparison of H-Infinity and Ensemble Kalman Filter for Estimating Motion of Middle Finger [22].	Both methods, even the Ensemble Kalman Filter and the H-Infinity, have accuracy above 98%

Source	Title	Key Findings
Katias *et al.*	Unscented Kalman Filter and H-Infinity for Travel Company Stock Price Estimation [23]	Unscented Kalman Filter reached good prediction with an accuracy error of less than 2.4%
Anshori *et al.*	Profitability estimation of XYZ company using H-infinity and Ensemble Kalman Filter [24].	The Ensemble Kalman Filter has better accuracy than H-Infinity with an error of about 5-8%.
Xu *et al.*	System Bias Correction of Short-Term Hub-Height Wind Forecasts Using the Kalman Filter [25]	The result is Kalman Filter successfully approaches 72-h forecasts, has decreased RMSE by 16% from 3.58 to 3.01 ms^{-1}, MAE by 14% from 2.71 to 2.34 ms^{-1}, bias from 0.22 to 0.19 ms^{-1}, and improved correlation from 0.58 to 0.66.
Qin, *et al.*	Trajectory Prediction Based on Long Short-Term Memory Network and Kalman Filter Using Hurricanes as an Example [26].	Better prediction using a combination of LSTM-KF than improved LSTM and simple LSTM.
Arora *et al.*	Prediction And Analysis Of COVID-19 Positive Cases Using Deep Learning Models: A Descriptive Case Study of India [27].	The result is an RNN with a bi-directional LSTM model that produced better predictions than a convolutional-LSTM.
Karya *et al.*	Estimation Of Crude Oil Price Using Unscented Kalman Filter [28].	This study resulted in an Unscented Kalman Filter reaching the best error rate of 2% and a Kalman Filter 8%.

2.2 Data Acquisition

The time-series data used in this study was obtained from the BPS East Java province (*https://jatim.bps.go.id*), having 3 columns and 114 rows with a time span of January 02, 2018, to December 18, 2020, which describes about condition of HIV/AIDS cases in East Java province. Before in-depth analysis, the data was cleaned of missing values, outliers, and other noise. This is all to simplify comprehension and improve the quality of the data to be tested, and to ease the decision of the machine learning model used. An overview of the research data and flowchart can be seen in Table 2 and Fig. 1 below.

Table 2. HIV Cases Dataset

Year	Name Of Regency/City	HIV Cases
02/01/2018	Pacitan	20
11/01/2018	Ponorogo	102
25/01/2018	Trenggalek	88
02/02/2018	Tulungagung	464
11/02/2018	Blitar	183
25/02/2018	Kediri	227
02/03/2018	Malang	382
11/03/2018	Lumajang	476
21/03/2018	Jember	792
29/03/2018	Banyuwangi	643
02/04/2018	Bondowoso	119
...	...	...
18/12/2020	Batu	153

After obtaining and exploring the data, the next step is to carry out the stages of the research flowchart one by one, as shown in Fig. 1 below.

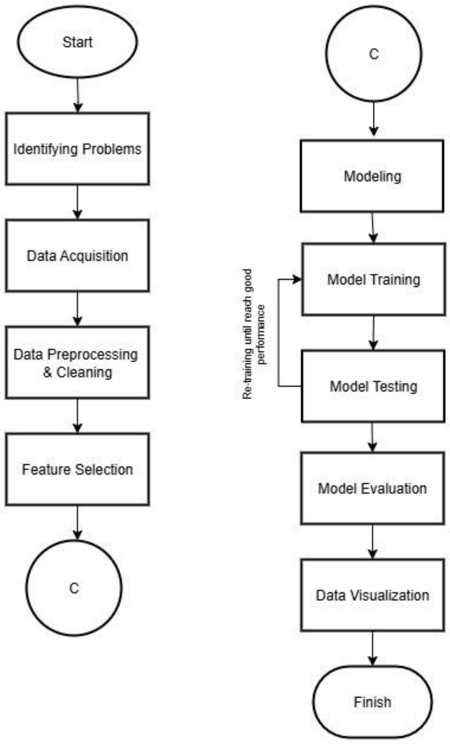

Figure 1. Research Flowchart

2.3 Exploratory Data Analysis

This research presents a case study on the prediction of HIV/AIDS cases in East Java using the data from BPS East Java Province. Based on the data obtained, there are three main variables, namely Date, Regency/City, and HIV cases. Variable HIV cases were used as the target variable/dependent. Below in Fig. 2 is the condition of the data based on the original source, and Table 3 is the table of central tendency.

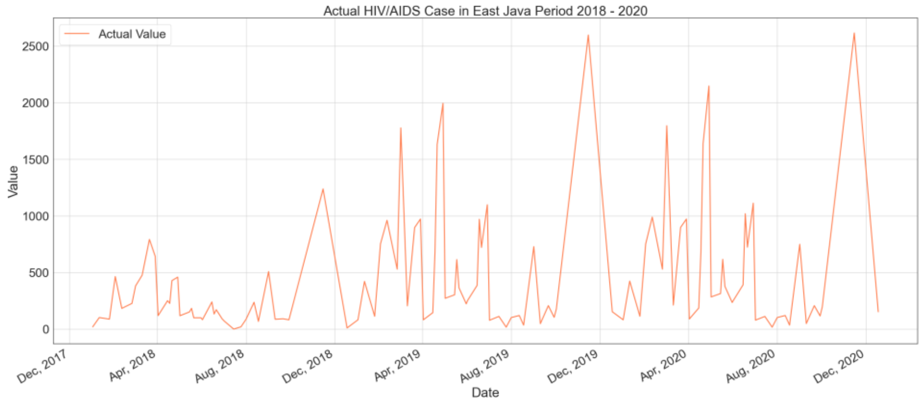

Figure 2. Plot of HIV AIDS Case from January 02, 2018, to December 18, 2020

Table 3. Measure of Central Tendency

	HIV CASES
mean	449.21
min	0
max	2615
std	537.55

After that, the data is refined by normalizing the values. The function of this normalization is to equalize the range of values between 0-1. Below in Eq. (1) is the function to normalize data, namely the Min Max Scaler.

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)}, \tag{1}$$

description:

x' : new values;
x : original values;
$\min(x)$: minimum values;
$\max(x)$: maximum values.

2.4 Kalman Filter Model

Kalman Filter, which was proposed by Rudolf Emil Kalman to complete the linear filtering problem in the field of aircraft and aviation, is one of the most significant and popular estimation methods [29]. The Kalman Filter is known as the optimal estimator that minimizes the mean squared error when the state and measurement dynamics are modeled as linear, and the process/measurement noise processed are modeled as white Gaussian [30]. Below in Eqs. (2)-(11) is the Kalman Filter algorithm.

System Model and Measurement

$$\boldsymbol{x_k} = \boldsymbol{F x_{k-1}} + \boldsymbol{B u_{k-1}} + \boldsymbol{w_{k-1}}, \tag{2}$$

$$\boldsymbol{z_k} = \boldsymbol{H x_k} + \boldsymbol{v_k}, \tag{3}$$

$$x_0 \sim N(\bar{x}_0, P_{x_0}); w_k \sim N(0, Q_k);\ v_k \sim N(0, R_k). \tag{4}$$

Initialization

$$\boldsymbol{\hat{x}_0} = \boldsymbol{\bar{x}_0}, \tag{5}$$

$$\boldsymbol{p_0} = \boldsymbol{p_{x_0}}. \tag{6}$$

Prediction stage

Estimation

$$\hat{x}_{k|k-1} = F_{\hat{x}_{k-1|k-1}} + B u_{k-1}. \tag{7}$$

Error covariance

$$P_{k|k-1} = F P_{k-1|k-1} F^T + Q. \tag{8}$$

Update stage

Kalman Gain

$$K_k = P_{k|k-1} H^T \left(H P_{k|k-1} H^T + R\right)^{-1}. \tag{9}$$

Estimation

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k\left(z_k - H_k \hat{x}_{k|k-1}\right). \tag{10}$$

Error covariance

$$P_{k|k} = [I - K_k H] P_{k|k-1}. \tag{11}$$

description:

x_k : state vector at time k;
F : state transition matrix;
$\boldsymbol{u_{k-1}}$: control input vector;
B : control input matrix;
$\boldsymbol{w_{k-1}}$: process noise (Q);
z_k : measurement vector;
$\boldsymbol{H}$: observation matrix;
v_k : measurement noise (R);
$\hat{x}_{k|k-1}$: a priori state estimate at step k;
$\hat{x}_{k|k}$: a posteriori state estimate at step k;

$\boldsymbol{P_{k|k-1}}$: a priori error covariance matrix;
$\boldsymbol{P_{k|k}}$: a posteriori error covariance matrix;
$P_{k|k}$: Kalman gain;
$z_k - H_k \hat{x}_{k|k-1}$: measurement innovation/residual.

2.5 Recurrent Neural Network Model (RNN)

A Recurrent Neural Network (RNN) is an improved form of neural networks that has internal memory, which makes RNN capable of handling long sequences [31]. In this study, the RNN type used is Simple RNN. Below in Eq. (12) is the function of the Recurrent Neural Network (RNN).

$$\boldsymbol{h_k} = \boldsymbol{\sigma_k}(\boldsymbol{U}\boldsymbol{h_{k-1}} + \boldsymbol{W}\boldsymbol{s_k} + \boldsymbol{b}), k = 0,1 \dots, N. \tag{12}$$

description:

k : discrete time index;
N : final finite horizon time;
$\boldsymbol{s_k}$: $m - d$ vector input sequence;
$\boldsymbol{\sigma_k}$: general nonlinear function (sigmoid/tanh).

2.6 Decision Tree (DT)

DT is a flow chart-like structure that uses a branching technique to define each possibility of results decisions [32]. Below in Eq. (13) are functions of the Decision Tree.

1. Estimate the entropy $E(S)$ value of the DS as written in Eq. (14)

$$E(S) = \sum_{i=1}^{m} -p_i log_2 p_i \tag{13}$$

where $E(S)$ = entropy of a collection of DS, m represents the number of classes in the system and p_i represents the number of instances that belong to class i.

2. Calculate the information gain for an attribute K, in a collection S, as expressed in Eq. (14), where $E(S)$ represents the entropy of the entire collection and S_u expressed the set of instances that have value u for attribute K.

$$G(S,K) = E(S) - \sum_{u \in values(K)} \frac{S_u}{S} E(S_u). \tag{14}$$

2.7 Evaluation Metrics

At the evaluation stage, the model trained and tested is calculated for accuracy based on the resulting error value. This study uses the Root Mean Square Error (RMSE) method to calculate the error value generated by the model. One of the main advantages of using RMSE is to provide higher weightage (as it contains a square) to larger errors [33] also MAE and MSE. The function of the Root Mean Square Error (RMSE) and the other methods are in Eqs. (15)-(17) as follows:

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |Y_i - \hat{Y}_i|, \tag{15}$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2, \tag{16}$$

$$RMSE = \sqrt{\frac{\sum(y_i - \hat{y}_i)\,2}{n}}. \tag{17}$$

description:

n : number of predictions;
Y_i : Observed value;
$\hat{Y}_i$: Predicted value.

3. RESULTS AND DISCUSSION

3.1 Hyperparameter Tuning and Data Splitting

The following Table 4 is a description of data splitting, and Table 5 - Table 7 are descriptions of parameter adjustments of the Kalman Filter, the Recurrent Neural Network, and the Decision Tree used to build a prediction model. In this study tried to implement accuracy testing from each method by a percentage of splitting data 80 of 20, also using static parameters to consider the limited amount of data. The variable to be used as the dependent variable is HIV cases.

Table 4. Data Splitting

Splitting Percentage	Number of Data Training	Number of Data Testing
80% : 20%	91	23

Table 5. Kalman Filter Model Hyperparameter

Parameter	Value
R value	0.03
Q value	0.03

Table 6. RNN Model Hyperparameter

Parameter	Value
Learning Rate	0.001
Stopping Condition	100

Table 7. Decision Tree Model Hyperparameter

Parameter	Value
Max depth	6
Min samples split	5

3.2 Result of Simulation

From the testing simulation conducted with the proportion of training data and testing data, the following results can be seen in Figs. 3 - 5.

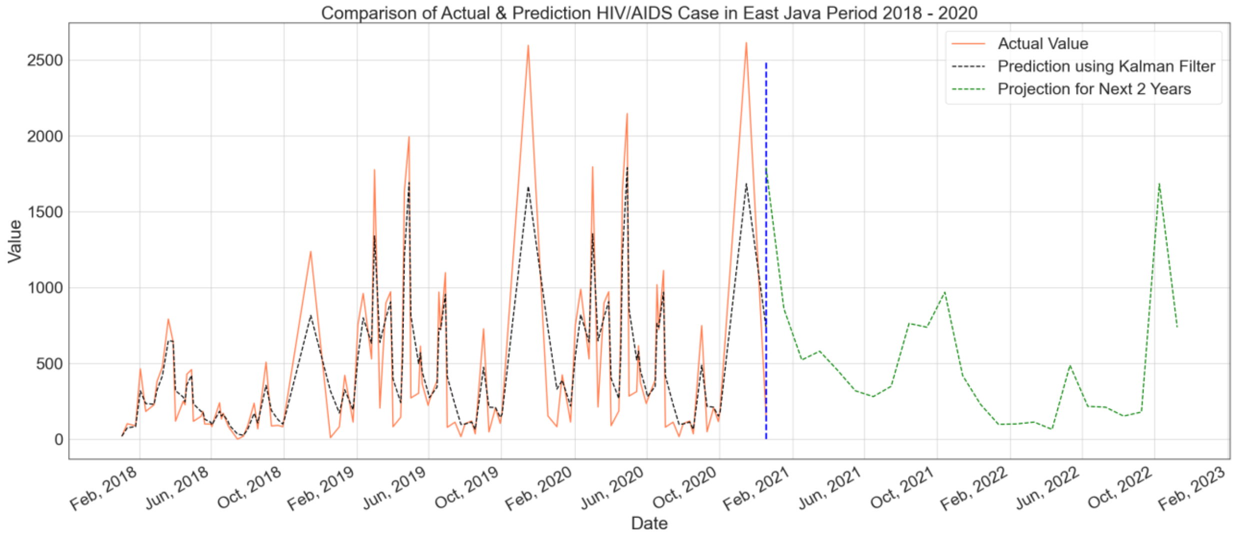

Figure 3. Simulation Plot of Kalman Filter using Q and R Value of 0.03

Fig. 3 shows the results of prediction using the Kalman Filter. The results achieved by this method are quite satisfactory. In the plot above, the predicted values are shown by the black line, which closely approximates the actual values. The green line represents the results of the Kalman Filter prediction projection for a period of 2 years. The prediction projection is done by taking the last 24 data points for training, which are then predicted again with the same hyperparameters. The results of the Kalman Filter method projection appear to be quite dynamic. In this simulation, the Kalman Filter method obtained an MAE value of 0.056, then an MSE value of 0.0078, and an RMSE of 0.0885, respectively.

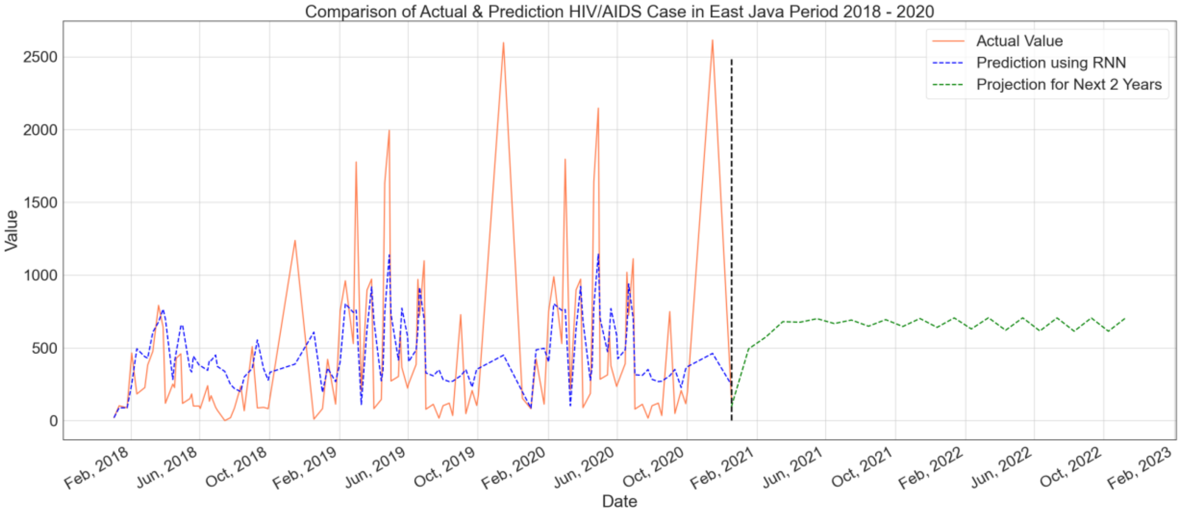

Figure 4. Simulation Plot of Recurrent Neural Network

Fig. 4 shows the results of prediction using a Recurrent Neural Network (RNN). The results achieved by this method are not quite satisfactory. In the plot above, the predicted values are shown by the blue line, which closely approximates the actual values. The green line represents the results of this method's prediction projection for a period of 2 years. The prediction projection is done by taking the last 24 data points for training, which are then predicted again with the same hyperparameters. The results of the RNN method projection appear to be quite flat. In this simulation, the RNN method obtained an MAE value of 0.123, then MSE value is 0.0395, and RMSE value of 0.198, respectively.

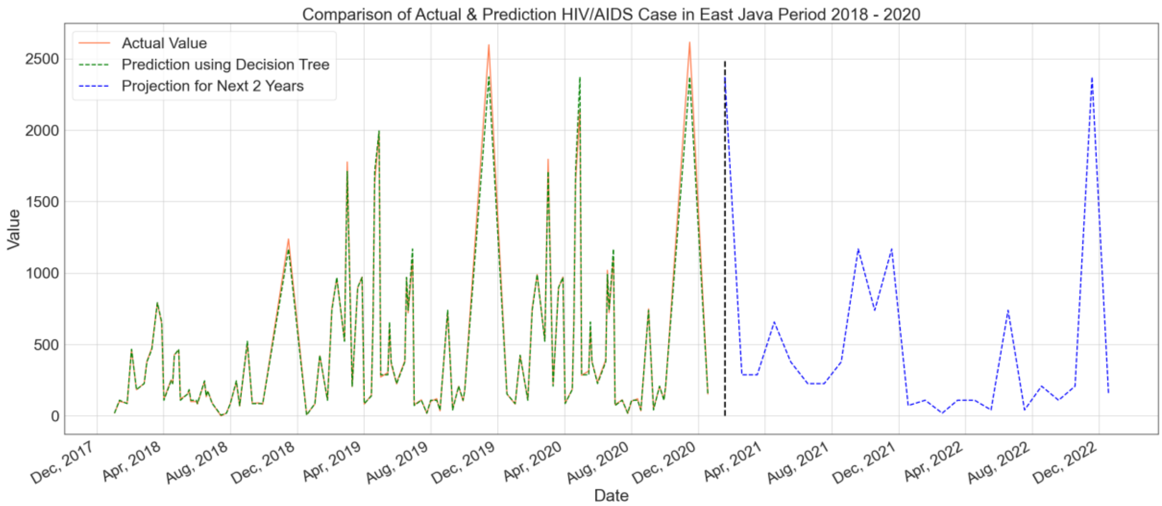

Figure 5. Simulation Plot of Decision Tree

Fig. 5 is the result of prediction using a Decision Tree. The results achieved by this method are quite satisfactory. In the plot above, the predicted values are shown by the green line, which closely approximates the actual values. The blue line represents the results of this method's prediction projection for a period of 2 years. The prediction projection is done by taking the last 24 data points for training, which are then predicted again with the same hyperparameters. The results of the Decision Tree method projection appear to be quite dynamic and able to follow the previous trend. In this simulation, the Decision Tree method obtained an MAE value of 0.0155, an MSE value of 0.0008, and an RMSE of 0.0287, respectively.

Based on the simulation results above, a recapitulation of the simulation results of the Kalman, Recurrent Neural Network (RNN), and Decision Tree can be seen in Table 8 below.

Table 8. Model Evaluation

Method	MAE	MSE	RMSE
Kalman Filter	0.056	0.0078	0.0885
RNN	0.123	0.0395	0.198
Decision Tree	0.0155	0.0008	0.0287

From Table 8, it can be shown that the Kalman Filter, the Recurrent Neural Network, and the Decision Tree produce satisfactory model evaluation values. By all simulations, the Decision Tree method successfully achieved the best evaluation with an MAE value of 0.0155, then the MSE value is 0.0008, and the RMSE is 0.0287.

4. CONCLUSION

Based on the results of the simulations conducted, it can be concluded that the results are as follows:

1. The Decision Tree method successfully achieved the best evaluation with an MAE value of 0.0155, then the MSE value is 0.0008, and the RMSE is 0.0287, respectively. This method can also capture fluctuating data conditions well, even when using a univariate basis.
2. The Kalman Filter method also achieved a satisfactory prediction result, which obtained an MAE value of 0.056, an MSE value of 0.0078, and an RMSE of 0.0885, respectively. As another note, in the simulation process, this method also has the advantage of shorter computational time than the other two methods, and also captures fluctuating data conditions well even when using a univariate basis.
3. In the current study, the results of implementing the RNN prediction method were not very satisfactory. The factor that concerns us is the lack of sufficient data, and it will be improved for the next study.

These results prove that Decision Tree and Kalman Filter methods provide good and consistent prediction results, also has fulfilled the objective of this research. Both methods can be used by health stakeholders to assist in planning and decision-making. In future studies, we recommend using other classical methods such as ARIMA or LSTM, which are subsets of deep learning methods and ensemble trees.

Author Contributions

Nurwijayanti: Conceptualization, Methodology, Writing-Original Draft, Software, Validation. Firman Yudianto: Data Curation, Resources, Draft Preparation. Panca Radono: Formal Analysis, Validation. Rizki Amalia Sinulingga: Software, Visualization. Mochammad Romli Arief: Validation, Writing-Review and Editing. Rahayu Budi Utami: Validation, Writing-Review and Editing. Hamzah Arof: Validation, Writing-Review and Editing. All authors discussed the results and approved the final version of the manuscript.

Funding Statement

This study was entirely self-supported and did not benefit from any form of external financial assistance. The authors affirm that no funding was obtained from governmental bodies, commercial enterprises, or non-profit organizations for the design and implementation of the research, the collection, analysis, and interpretation of data, the preparation of the manuscript, or the decision to submit the article for publication. All expenses associated with the research process and manuscript preparation were borne solely by the authors.

Acknowledgment

The authors gratefully acknowledge for colleagues and peers for their valuable insights, constructive comments, and encouragement during the research process and manuscript preparation. Although this study did not receive external funding, the intellectual and moral support from the academic community has been invaluable in enhancing the quality and rigor of the work presented herein.

Declarations

The authors explicitly state that there are no conflicts of interest that could be perceived as influencing the objectivity, integrity, or outcomes of this study. The research was conducted independently without any financial, professional, or personal relationships that might be considered relevant to the content of this manuscript. All authors confirm that they have no affiliations, financial involvements, or other connections with organizations or entities that could inappropriately bias or compromise the work presented herein.

Declarations of Generative AI and AI-assisted Technologies

Generative AI tools (e.g., ChatGPT) was used solely for language refinement, including grammar, spelling, and clarity. The scientific content, analysis, interpretation, and conclusions were developed entirely by the authors. All final text was reviewed and approved by the authors.

REFERENCES

[1] N. S. Rahmi, S. Astutik, A. B. Astuti, A. R. Muhammad, U. Maisaroh, and S. Handayani, "GENERALIZED CONFIRMATORY FACTOR ANALYSIS FOR KNOWING IMPACT OF KNOWLEDGE, ATTITUDES, AND BEHAVIORAL FACTORS HIV/AIDS IN INDONESIA," BAREKENG JURNAL ILMU MATEMATIKA DAN TERAPAN, vol. 17, no. 2, pp. 0695-0706, Jun. 2023. doi: https://doi.org/10.30598/barekengvol17iss2pp0695-0706.

[2] T. D. Chandra and G. I. Permata, "MODELING HIV/AIDS USING SHAT MODEL," BAREKENG JURNAL ILMU MATEMATIKA DAN TERAPAN, vol. 17, no. 2, pp. 0745–0756, Jun. 2023. doi: https://doi.org/10.30598/barekengvol17iss2pp0745-0756.

[3] S. Davlidova, Z. Haley-Johnson, K. Nyhan, A. Farooq, S. H. Vermund, and S. Ali, "PREVALENCE OF HIV, HCV AND HBV IN CENTRAL ASIA AND THE CAUCASUS: A SYSTEMATIC REVIEW," *International Journal of Infectious Diseases*, vol. 104, pp. 510–525, Mar. 2021. doi: https://doi.org/10.1016/J.IJID.2020.12.068.

[4] R. Janamnuaysook *et al.*, "DEMEDICALISATION OF HIV INTERVENTIONS TO END HIV IN THE ASIA-PACIFIC," *Sexual Health*, vol. 18, no. 1, pp. 13–20, Feb. 2021. doi: https://doi.org/10.1071/SH20172.

[5] R. A. I. Sadarang, "PREVALENCE AND FACTORS AFFECTING DISCRIMINATION TOWARDS PEOPLE LIVING WITH HIV/AIDS IN INDONESIA," *Journal of Preventive Medicine and Public Health*, vol. 55, no. 2, pp. 205–212, Mar. 2022. doi: https://doi.org/10.3961/jpmph.21.502.

[6] C. Kirana, R. Tosepu, and D. S. Effendy, "HIV/AIDS POSITIVE CASES BASED ON BASIC HEALTH RESEARCH DATA IN 2019, INDONESIA," *KnE Life Sciences*, Sep. 2022. doi: https://doi.org/10.18502/kls.v0i0.11781.

[7] "MODELLING THE NUMBER OF HIV CASES IN INDONESIA USING NEGATIVE BINOMIAL REGRESSION BASED ON LEAST SQUARE SPLINE ESTIMATOR," *Communications in mathematical biology and neuroscience*, Jan. 2024. doi: https://doi.org/10.28919/cmbn/8582.

[8] M. D. Kurniasari, A. D. Huruta, H. T. Tsai, and C.-W. Lee, "FORECASTING FUTURE HIV INFECTION CASES: EVIDENCE FROM INDONESIA," *Social Work in Public Health*, vol. 36, no. 1, pp. 12–25, Dec. 2020. doi: https://doi.org/10.1080/19371918.2020.1851332.

[9] R. Fitriani and A. Lutfi, "KNOWLEDGE OF HIV TRANSMISSION AND FACTORS RELATED TO THE INCIDENCE OF HIV/AIDS IN ADOLESCENTS IN INDONESIA," *Indian Journal of Forensic Medicine & Toxicology*, Jan. 2021. doi: https://doi.org/10.37506/ijfmt.v15i1.13618.

[10] D. M. S. K. Dewi, J. D. E. Sari, M. Z. Fatah, and E. Astutik, "STIGMA AND DISCRIMINATION AGAINST PEOPLE LIVING WITH HIV AND AIDS IN BANYUWANGI, EAST JAVA, INDONESIA," *Proceedings of the 4th International Symposium on Health Research (ISHR 2019)*, Jan. 2020. doi: https://doi.org/10.2991/ahsr.k.200215.030.

[11] Z. M. Kusumaadhi, N. Farhanah, and M. A. U. Sofro, "RISK FACTORS FOR MORTALITY AMONG HIV/AIDS PATIENTS," *Diponegoro International Medical Journal*, vol. 2, no. 1, pp. 20–19, Mar. 2021. doi: https://doi.org/10.14710/dimj.v2i1.9667.

[12] S. G. Aribe Jr, B. D. Gerardo, and R. P. Medina, "TIME SERIES FORECASTING OF HIV/AIDS IN THE PHILIPPINES USING DEEP LEARNING: DOES COVID-19 EPIDEMIC MATTER?," *International Journal of Emerging Technology and Advanced Engineering*, vol. 12, no. 9, pp. 144–157, Sep. 2022. doi: https://doi.org/10.46338/ijetae0922_15.

[13] C. K. Mutai, P. E. McSharry, I. Ngaruye, and E. Musabanganji, "USE OF MACHINE LEARNING TECHNIQUES TO IDENTIFY HIV PREDICTORS FOR SCREENING IN SUB-SAHARAN AFRICA," *BMC Medical Research Methodology*, vol. 21, no. 1, Jul. 2021. doi: https://doi.org/10.1186/s12874-021-01346-2.

[14] T. Herlambang, A. Suryowinoto, D. Adrianto, D. Rahmalia, and H. Nurhadi, "DIVING MOTION ESTIMATION OF REMOTELY OPERATED VEHICLE USING ENSEMBLE KALMAN FILTER AND H-INFINITY", *BAREKENG: J. Math. & App.*, vol. 17, no. 1, pp. 0095-0100, Apr. 2023. https://doi.org/10.30598/barekengvol17iss1pp0095-0100

[15] Muhith, T. Herlambang, D. Rahmalia, and D. F. Karya, "ESTIMATION OF PACKED RED CELLS (PRC) IN BOJONEGORO BLOOD BANK USING MODIFIED KALMAN FILTER," *Journal of Physics Conference Series*, vol. 2157, no. 1, p. 012023, Jan. 2022. doi: https://doi.org/10.1088/1742-6596/2157/1/012023.

[16] H. Hewamalage, C. Bergmeir, and K. Bandara, "RECURRENT NEURAL NETWORKS FOR TIME SERIES FORECASTING: CURRENT STATUS AND FUTURE DIRECTIONS," International Journal of Forecasting, vol. 37, no. 1, pp. 388–427, Aug. 2020. doi: https://doi.org/10.1016/j.ijforecast.2020.06.008.

[17] K. E. ArunKumar, D. V. Kalaga, C. S. Kumar, M. Kawaji, and T. M. Brenza, "FORECASTING OF COVID-19 USING DEEP LAYER RECURRENT NEURAL NETWORKS (RNNS) WITH GATED RECURRENT UNITS (GRUS) AND LONG SHORT-TERM MEMORY (LSTM) CELLS," *Chaos Solitons & Fractals*, p. 110861, Mar. 2021. doi: https://doi.org/10.1016/j.chaos.2021.110861.

[18] M. N. Fekri, H. Patel, K. Grolinger, and V. Sharma, "DEEP LEARNING FOR LOAD FORECASTING WITH SMART METER DATA: ONLINE ADAPTIVE RECURRENT NEURAL NETWORK," *Applied Energy*, vol. 282, p. 116177, Nov. 2020. doi: https://doi.org/10.1016/j.apenergy.2020.116177.

[19] Ahmadi, M. Nabipour, B. Mohammadi-Ivatloo, A. M. Amani, S. Rho, and Md. J. Piran, "LONG-TERM WIND POWER FORECASTING USING TREE-BASED LEARNING ALGORITHMS," *IEEE Access*, vol. 8, pp. 151511–151522, Jan. 2020. doi: https://doi.org/10.1109/access.2020.3017442.

[20] M. Bansal, A. Goyal, and A. Choudhary, "A COMPARATIVE ANALYSIS OF K-NEAREST NEIGHBOR, GENETIC, SUPPORT VECTOR MACHINE, DECISION TREE, AND LONG SHORT-TERM MEMORY ALGORITHMS IN MACHINE LEARNING," *Decision Analytics Journal*, vol. 3, p. 100071, May 2022. doi: https://doi.org/10.1016/j.dajour.2022.100071.

[21] Setiawan, F. Setivani, and T. Mahatma, "PERFORMANCE COMPARISON OF DECISION TREE AND LOGISTIC REGRESSION METHODS FOR CLASSIFICATION OF SNP GENETIC DATA," BAREKENG JURNAL ILMU MATEMATIKA DAN TERAPAN, vol. 18, no. 1, pp. 0403-0412, Mar. 2024. doi: https://doi.org/10.30598/barekengvol18iss1pp0403-0412.

[22] T. Herlambang, H. Nurhadi, F. Susanto, and B. Tomasouw, "COMPARISON OF H-INFINITY AND ENSEMBLE KALMAN FILTER FOR ESTIMATING MOTION OF MIDDLE FINGER", *BAREKENG: J. Math. & App.*, vol. 17, no. 4, pp. 2433-2442, Dec. 2023. doi: https://doi.org/10.30598/barekengvol17iss4pp2433-2442

[23] P. Katias, I. Susanto, T. Herlambang, and M. Anshori, "UNSCENTED KALMAN FILTER AND H-INFINITY FOR TRAVEL COMPANY STOCK PRICE ESTIMATION", *PIJMath*, vol. 1, no. 2, pp. 81-88, Nov. 2022. doi: https://doi.org/10.30598/pijmathvol1iss2pp81-88

[24] M. Y. Anshori, T. Herlambang, P. Katias, F. A. Susanto, and R. R. Rasyid, "PROFITABILITY ESTIMATION OF XYZ COMPANY USING H-INFINITY AND ENSEMBLE KALMAN FILTER," *Journal of Physics Conference Series*, vol. 2157, no. 1, p. 012027, Jan. 2022. doi: https://doi.org/10.1088/1742-6596/2157/1/012027.

[25] J. Xu, Z. Xiao, Z. Lin, and M. Li, "SYSTEM BIAS CORRECTION OF SHORT-TERM HUB-HEIGHT WIND FORECASTS USING THE KALMAN FILTER," *Protection and Control of Modern Power Systems*, vol. 6, no. 1, Nov. 2021. doi: https://doi.org/10.1186/s41601-021-00214-x.

[26] W. Qin, J. Tang, C. Lu, and S. Lao, "TRAJECTORY PREDICTION BASED ON LONG SHORT-TERM MEMORY NETWORK AND KALMAN FILTER USING HURRICANES AS AN EXAMPLE," *Computational Geosciences*, vol. 25, no. 3, pp. 1005–1023, Mar. 2021. doi: https://doi.org/10.1007/s10596-021-10037-2.

[27] P. Arora, H. Kumar, and B. K. Panigrahi, "PREDICTION AND ANALYSIS OF COVID-19 POSITIVE CASES USING DEEP LEARNING MODELS: A DESCRIPTIVE CASE STUDY OF INDIA," *Chaos Solitons & Fractals*, vol. 139, p. 110017, Jun. 2020. doi: https://doi.org/10.1016/j.chaos.2020.110017.

[28] D. F. Karya, M. Y. Anshori, R. Rizqina, P. Katias, A. Muhith, and T. Herlambang, "ESTIMATION OF CRUDE OIL PRICE USING UNSCENTED KALMAN FILTER," *Journal of Physics Conference Series*, vol. 1538, p. 012049, May 2020. doi: https://doi.org/10.1088/1742-6596/1538/1/012049.

[29] S. Feng, X. Li, S. Zhang, Z. Jian, H. Duan, and Z. Wang, "A REVIEW: STATE ESTIMATION BASED ON HYBRID MODELS OF KALMAN FILTER AND NEURAL NETWORK," *Systems Science & Control Engineering*, vol. 11, no. 1, Feb. 2023. doi: https://doi.org/10.1080/21642583.2023.2173682.

[30] R. V. Vitali, R. S. McGinnis, and N. C. Perkins, "ROBUST ERROR-STATE KALMAN FILTER FOR ESTIMATING IMU ORIENTATION," *IEEE Sensors Journal*, vol. 21, no. 3, pp. 3561–3569, Sep. 2020. doi: https://doi.org/10.1109/jsen.2020.3026895.

[31] P. Dey *et al.*, "COMPARATIVE ANALYSIS OF RECURRENT NEURAL NETWORKS IN STOCK PRICE PREDICTION FOR DIFFERENT FREQUENCY DOMAINS," *Algorithms*, vol. 14, no. 8, p. 251, Aug. 2021. doi: https://doi.org/10.3390/A14080251.

[32] K. Nti, A. F. Adekoya, and B. A. Weyori, "A COMPREHENSIVE EVALUATION OF ENSEMBLE LEARNING FOR STOCK-MARKET PREDICTION," *Journal of Big Data*, vol. 7, no. 1, Mar. 2020. doi: https://doi.org/10.1186/s40537-020-00299-5.

[33] P. K. Jena, S. M. Rahaman, P. K. D. Mohapatra, D. P. Barik, and D. S. Patra, "SURFACE WATER QUALITY ASSESSMENT BY RANDOM FOREST," *Water Practice & Technology*, vol. 18, no. 1, pp. 201–214, Dec. 2022. doi: https://doi.org/10.2166/wpt.2022.156.

[34] "PANDAS DOCUMENTATION — PANDAS 2.3.2 DOCUMENTATION."https://pandas.pydata.org/docs/index.html

[35] C. R. Harris *et al.*, "ARRAY PROGRAMMING WITH NUMPY," *Nature*, vol. 585, no. 7825, pp. 357–362, Sep. 2020. doi: https://doi.org/10.1038/s41586-020-2649-2.

[36] J. D. Hunter, "MATPLOTLIB: A 2D GRAPHICS ENVIRONMENT," *Computing in Science & Engineering*, vol. 9, no. 3, pp. 90–95, Jan. 2007. doi: https://doi.org/10.1109/mcse.2007.55.

[37] M. Waskom, "SEABORN: STATISTICAL DATA VISUALIZATION," *The Journal of Open Source Software*, vol. 6, no. 60, p. 3021, Apr. 2021. doi: https://doi.org/10.21105/joss.03021.

[38] Scikit-learn Developers, "Scikit-learn: Machine Learning in Python (Version 1.7.1) Documentation," 2025. [Online]. Available: https://scikit-learn.org/stable/index.html.