IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 13, No. 2, June 2024, pp. 2036~2048
ISSN: 2252-8938, DOI: 10.11591/ijai.v13.i2.pp2036-2048

A hybrid deep learning optimization for predicting the spread of a new emerging infectious disease

Faulinda Ely Nastiti[1,2], Shahrulniza Musa[1,4], Eiad Yafi[3]
[1]Malaysian Institute of Information Technology, Universiti Kuala Lumpur, Kuala Lumpur, Malaysia
[2]Department of Information System, Faculty of Computer Science, Universitas Duta Bangsa Surakarta, Surakarta, Indonesia
[3]School of Computer Science, University of Technology Sydney, Sydney, Australia
[4]UniKL-LR Univ Joint ICT Laboratory (KLR-JIL), Universiti Kuala Lumpur, Malaysia-La Rochelle University, La Rochelle, France

Article Info

Article history:

Received Jul 31, 2023
Revised Oct 9, 2023
Accepted Dec 20, 2023

Keywords:

COVID-19
Grey wolf optimization
Long-short-term memory
Particle swarm optimization
Prediction

ABSTRACT

In this study, a novel approach geared toward predicting the estimated number of coronavirus disease (COVID-19) cases was developed. Combining long short-term memory (LSTM) neural networks with particle swarm optimization (PSO) along with grey wolf optimization (GWO) employ hybrid optimization algorithm techniques. This investigation utilizes COVID-19 original data from the Ministry of Health of Indonesia, period 2020-2021. The developed LSTM-PSO-GWO hybrid optimization algorithm can improve the performance and accuracy of predicting the spread of the COVID-19 virus in Indonesia. In initiating LSTM initial weights with weaknesses, using the hybrid optimization algorithm helps overcome these problems and improve model performance. The results of this study suggest that the LSTM-PSO-GWO model can be utilized as an effective and reliable predictive tool to gauge the COVID-19 virus's spread in Indonesia.

Corresponding Author:

Shahrulniza Musa
Malaysian Institute of Information Technology, University Kuala Lumpur Kuala Lumpur
St. Sultan Ismail, Bandar Wawasan, Kuala Lumpur 50250, Malaysia
Email: shahrulniza@unikl.edu.my

1. INTRODUCTION

Infectious diseases categorized as re-emerging are Polio, Ebola, and Lassa fever [1]. Meanwhile, the Avian Influenza Virus type H5N1, H9N2, H7N9, H5N6, severe acute respiratory syndrome (SARS), and Middle East respiratory syndrome (MERS), emerged as new emerging diseases [2], [3]. Coronavirus disease (COVID-19) was identified as a new emerging infectious disease, evolved from a human-infecting enzootic bat virus [4]. COVID-19 is quickly spreading [5] via contact with an infected item, the splash of bodily fluids from an infected patient, and even the COVID-19 virus can survive and infect via airborne transmission [6]–[8].

On April 1, 2021, the Health Ministry of the Republic of Indonesia announced that 1,547,854 individuals had tested positive for COVID-19. COVID-19 was a severe public health issue that produced injuries, deaths, and broad socioeconomic consequences [9]. Therefore, predicting the positive, mortality, and recovery rates must be enhanced to prevent the virus from spreading uncontrollably. The prediction can help local governments make timely decisions and policies, such as preparing medical facilities and deciding on border closures or lockdowns [10]–[13].

Several studies in the field of machine learning have been conducted to predict the COVID-19 spreading cases using a single method such as artificial neural network (ANN) [14], multilayer perception (MLP) [15], random forest [16], and Gaussian Nave Bayes [17]. Unfortunately, the construction of several individual algorithms often results in a decrease in model performance [18] as more and more parameters must be correctly optimized [19]. This weakness in the accuracy of predictions occurs because the data contains characteristics that vary between countries, such as non-seasonal, non-linear, and multivariate [20], [21].

Non-seasonal, non-linear, and multivariate COVID-19 data problems can be resolved using the time series model, namely the long short-term memory (LSTM) model [22]. The LSTM model is a deep learning model, more widely used due to its vanilla LSTM has comparable performance, is easier to understand, and is often better than other LSTM fields [23]. According to some findings [24]–[26] in the research of predicting the spreading of COVID-19, the LSTM model has a greater level of accuracy when compared to linear and non-linear regression models as well as support vector machines (SVMs), K-nearest neighbors (KNN), and autoregressive integrated moving average (ARIMA). Although vanilla LSTM is imperative for deep learning time series models, it is insensitive to random weight initializations [27], [28].

Particle swarm optimization (PSO) searches for the optimal value of the objective function by relocating a set of particles within a specified search space [29]. The results of several studies [30], [31] demonstrate that the utilization of PSO for the initial weighting problem in LSTM can enhance the accuracy of the time series prediction model, as well as expedite its convergence speed. Nevertheless, when the convergence rate is excessively rapid, PSO may encounter a predicament where it becomes trapped within a local optimum. To attain optimal outcomes in the training of the LSTM model, it is crucial to integrate diverse optimization techniques and meticulously fine-tune the model's parameters.

Similar to the PSO algorithm, the grey wolf optimization (GWO) algorithm is among the multitude of heuristic optimization algorithms [32] that have garnered immense popularity in the field. GWO demonstrates adaptability and applies to a wide variety of optimization tasks and problem domains [33]. GWOs enhanced exploratory characteristics can effectively mitigate the potential disadvantage of becoming confined in a local optimum that PSO may encounter.

This study proposes a hybrid optimization model with LSTM-PSO-GWO to establish initial weights and parameters to overcome LSTM problems, achieve a lower loss, and increase performance in predicting the spread of the COVID-19 pandemic virus. Due to the ever-changing complexity and dynamics of virus transmission, a more precise and adaptable predictive method is required to aid in decision-making and the development of appropriate preventative measures. A novel hybrid optimization model with LSTM-PSO-GWO as a better predictive model is crucial to efforts to combat this COVID-19 pandemic.

2. METHOD

2.1. Data-driven method

This research enhances a hybrid deep learning optimization model with the TensorFlow framework and runs it on Google Collab. LSTM module is incorporated with the PSO and GWO optimization algorithms in the proposed architecture in Figure 1. LSTM extracts temporal patterns from COVID-19 spread data, while PSO-GWO is employed to optimize the LSTM model's initial weight and parameters.

2.1. Data collecting

This study obtained original and direct data on the spread of COVID-19 from the data and information center of the Ministry of Health of the Republic of Indonesia because this data is the most authoritative and reliable source. The reason is to ensure the accuracy and up-to-date information regarding the distribution of daily cases, number of cases, recovery rate, deaths related to COVID-19, and other variables/features. Thus, using original data from the Ministry of Health provides a solid basis for the reliability and validity of our study findings. The spread of COVID-19 infection was obtained on December 20, 2021. The COVID distribution data has 80 features with a vulnerable time from February 3, 2020, to December 19, 2021.

2.3. Data preprocessing

The preprocessing was done in Figure 2, by cleaning the data, overcoming missing values, normalizing, and feature selection. The feature selection phase aims to enhance the model's reliability during training. The features used in each model are obtained by considering the correlation value of the target features; recovery, died, and positive. The utilized features have a correlation value greater than 0.80 or 80 percent of the objective feature. Each feature correlates differently with each target feature. Each target feature receives a distinct quantity of additional features after filtering.

On the target 'recovery', 38 additional features are obtained. On the target 'death', 39 additional features are obtained. In addition, one additional feature is acquired for the 'positive' target. The feature selection results can be observed in Table 1. The next step is splitting the dataset into three datasets: 60% training dataset, 20% validation dataset, and 20% testing dataset.

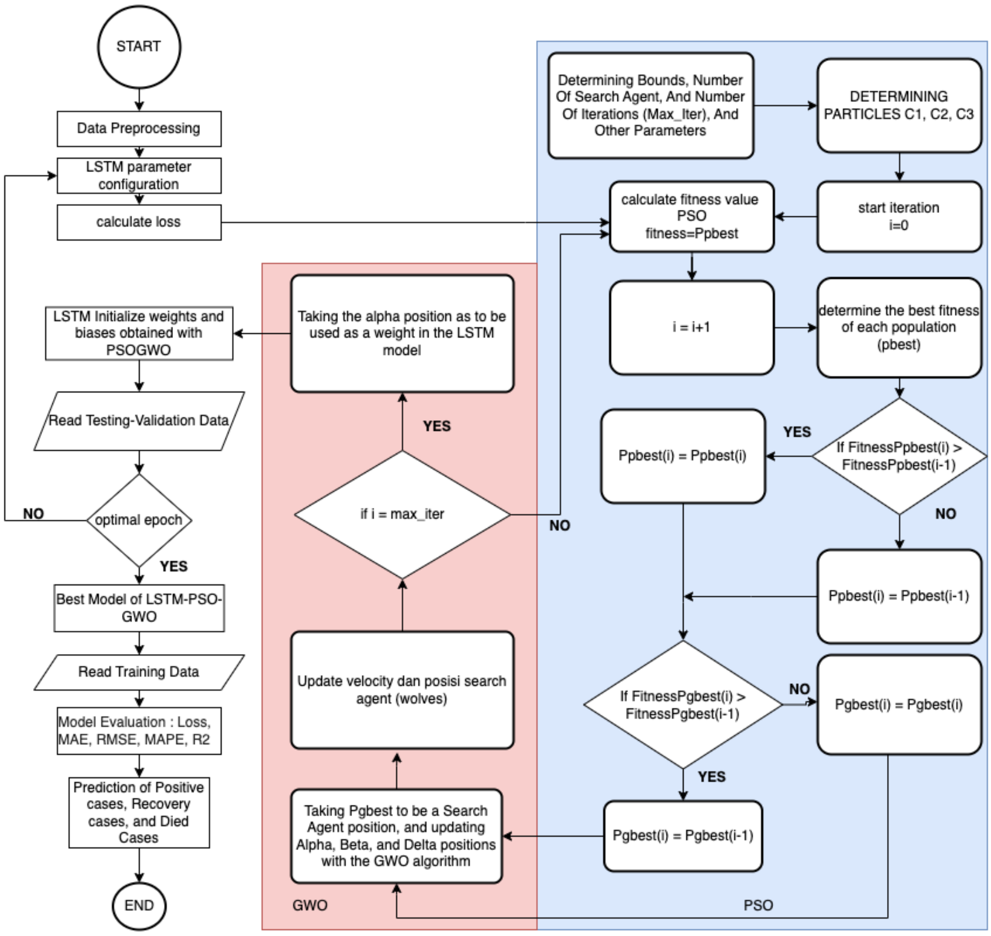


Figure 1. A hybrid deep learning optimization model LSTM-PSO-GWO proposed model

Figure 2. A hybrid deep learning optimization model LSTM-PSO-GWO preprocessing steps

Table 1. Feature's selection

Features target	Features selection
Recovery	'Total cases', 'Deaths', 'Number of specimens tested (since April 1)', 'Number of individuals tested', 'Negative', 'Specimens', 'Individuals tested', 'Individuals tested (Antigen)', 'Number of tests per million population', 'Average number of specimens tested (7-day average)', 'Average number of individuals tested (7-day average)', The terms 'First dose', 'Second dose', 'Third dose', 'First dose (%)', 'Second dose (%)', 'First dose (daily)', 'Second dose (daily)', 'Daily dose', 'First dose (weekly)', 'Second dose (weekly)', and 'Average daily dose (weekly)' are used in this context. The first dose (healthcare workers), the second dose (healthcare workers), the third dose (healthcare workers), the first dose (healthcare workers) percentage, the second dose (healthcare workers) percentage, the first dose (public service workers), the second dose (public service workers), the first dose (public service workers) percentage, the second dose (public service workers) percentage, the first dose (elderly), the second dose (elderly), the first dose (elderly) percentage, the second dose (elderly) percentage, the first dose (general population), the first dose (Vakgor), the second dose (Vakgor).
Death	'Total number of cases', 'Recovered', 'Number of specimens examined (since April 1)', 'Number of individuals examined'] 'Negative', 'Specimen', 'Tested individuals', 'Tested individuals (Antigen)', 'Number of tests per million population', 'Number of specimens examined (average over 7 days)', 'Number of individuals examined (average over 7 days)' The terms used in the text include 'First dose', 'Second dose', 'Third dose', 'First dose (%)', 'Second dose (%)', 'First dose (daily)', 'Second dose (daily)', 'Daily dose', 'First dose (weekly)', 'Second dose (weekly)', 'Average daily dose (weekly)', 'First dose (healthcare workers)', 'Second dose (healthcare workers)', 'Third dose (healthcare workers)', 'First dose (healthcare workers) (%)', 'Second dose (healthcare workers) (%)', 'First dose (public servants)', 'Second dose (public servants)', 'First dose (public servants) (%)', 'Second dose (public servants) (%)', 'First dose (elderly)', 'Second dose (elderly)', 'First dose (elderly) (%)', 'Second dose (elderly) (%)', 'First dose (general public)', 'Second dose (general public)', 'First dose (Vakgor)', 'Second dose (Vakgor)'.
Positive	'The recovery rate (closed cases)'

2.4. Long short-term memory configuration–hyperparameter

Within the scope of this inquiry, it is imperative to recognize the importance of hyperparameters, as they have the potential to exert a significant impact on the effectiveness of the optimization model and its convergence rate. In order to get the best results in the experiment, a predetermined set of essential hyperparameters has been established as outlined: the experimental parameters were a time length of 100 units, a batch size of 32, and a configuration of 32 hidden units. Moreover, in order to improve the optimization of the parameters in the LSTM model, it is essential to appropriately adjust many hyperparameters, including the combinations of gradian boosting and the PSO-GWO method.

Initially, the utilization of the 'get_shape()' function enables the acquisition of the model weights' shape, which holds significant importance as a fundamental hyperparameter in LSTM structures and several other techniques. In addition, the parameter search space was defined with upper and lower bounds of 1.0 and -1.0, respectively. The significance of this particular variable is in its influence on the performance of the model, as hyperparameters encompass a spectrum of values that can impact its overall effectiveness.

Once the hyperparameters have been established, the subsequent pivotal action is executing the 'optimizer. optimize()' function in order to optimize the model. The aforementioned function plays a crucial role in the initiation and supervision of the optimization process. The objective is to identify hyperparameter configurations that can yield optimal outcomes during the process of model training. Throughout the course of this iteration, the optimization algorithm will methodically investigate different permutations of hyperparameters in order to get optimal performance on the model presently undergoing training.

2.5. Hybrid modelling, training, and validation

This study presents a novel hybrid optimization model that has promise to make a significant contribution to the efforts in combating the COVID-19 pandemic. The main aim of this research is to improve the accuracy and effectiveness of the LSTM model by utilizing a hybrid optimization technique. The use of this technique will be employed on COVID-19 data, which displays complex characteristics like time series, non-linearity, and numerous variables. It is expected that the suggested model will have the capacity to predict the number of confirmed cases, fatality rates, and recovery rates for the upcoming seven-day period.

2.5.1. Long short-term memory architecture

Figure 3 depicts the architectural design of a LSTM neural network, which belongs to the category of recurrent neural networks renowned for their proficiency in processing sequential input. The graphic provides a clear depiction of the layers comprising the LSTM model, as well as the interconnections between the memory cells and the gates responsible for regulating the information flow from previous time steps. The presented visualization facilitates a comprehensive comprehension of the operations performed by LSTM models in tasks that involve sequential data.

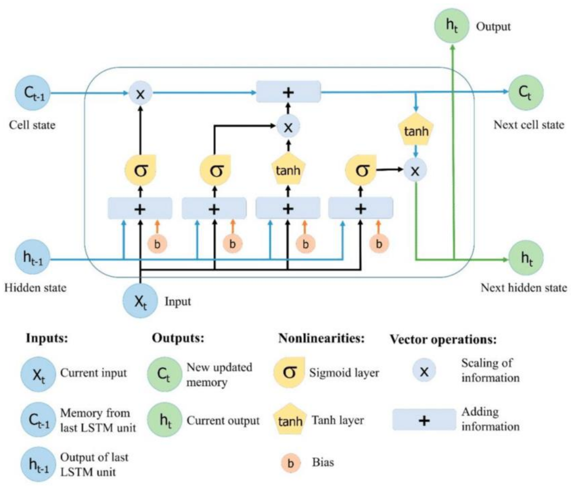


Figure 3. The LSTM neural network structure [34]

LSTM neural networks integrate timing into the network's structure, allowing them to be more flexible when applied to the analysis of time series data [26]. LSTM is used to process complex patterns datasets to produce accurate predictions on COVID-19 time series data: i) the network handles data from three different inputs, ii) X_t denotes the input for the current time step, iii) h_{t_1} is the output from the prior LSTM unit, and iv)

C_{t_1} is the "memory" of the prior unit. C_{t-1} input is the most crucial one. Regarding outputs, the current network's output is denoted by h_t. The memory of the current unit is denoted by the letter C_t [35].

The first stage in building an LSTM network is removing unnecessary information from the cell. The sigmoid function uses the output of the last LSTM unit (h_{t-1}) at time t−1 and the current input (X_t) at t to identify and exclude data. The sigmoid function also decides which old output should be removed. The forget gate (ft) is a vector with 0 to 1 value corresponding to each number in the cell state, C_{t-1}.

$$ft = \sigma(W_f[h_{t-1}, X_t] + b_f \tag{1}$$

The forget gate is represented by weight matrices (W_f) and bias (b_f), while the sigmoid function (σ) is utilized to determine the activation of the forget gate.

Next, the cell state is updated, and new input (X_t) information is stored. This phase has sigmoid and tanh layers. First, the sigmoid layer decides whether newly acquired data should be updated or ignored (0 or 1), and then the tanh function weights values that have passed by (1 to 1). Multiplying the two values updates the cell state. This new memory is added to C_{t1}, assembling C_t.

$$i_t = \sigma(W_i[h_{t-1}, X_t] + b_i) \tag{2}$$

$$N_t = tanh\,(W_n[h_{t-1}, X_t] + b_n) \tag{3}$$

$$C_t = C_{t-1}\, ft + N_t i_t) \tag{4}$$

At time t–1 and t, $C_{t-1,}$ and C_t represent the cell states, while the weight matrices (W) and bias (b) are associated with the cell state.

The final output values (h_t) are filtered from the output cell state (O_t). A sigmoid layer selects cell state components for output. Next, the sigmoid gate output (O_t) is multiplied by the new values formed by the tanh layer from the cell state (C_t), ranging from −1 to 1. W_o and b_o represent the output gate's weight matrix and bias.

$$O_t = \sigma(W_o[h_{t-1}, X_t] + b_o) \tag{5}$$

$$h_t = O_t \tanh(C_t) \tag{6}$$

2.5.2. Particle swam optimization algorithm

PSO is a population-based optimization technique inspired by the foraging behaviors of insect and bird colonies. PSO has a decent convergence rate compared to other population optimization techniques by searching in parallel and quickly approaching the optimal point [36], [37]. PSO illustration as shown in Figure 4.

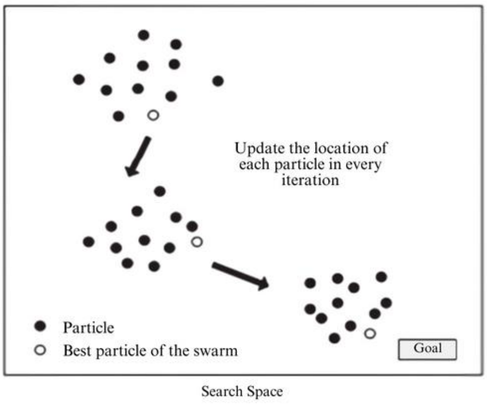


Figure 4. A swarm that follows the best particle to the destination [38]

Several particles are utilized to produce the swarm. Each particle represents a potential solution. The candidate solutions coexist and collaborate concurrently. Each particle in the swarm searches the search region for the optimal landing site. Hence, the search region represents the set of potential solutions, and the group (or swarm) of flying particles depicts the solutions as they evolve.

Throughout the generations (iterations), each particle maintains a record of both its own best solution (optimum) and the best solution (optimum) of the swarm. Then, it updates the flight speed (velocity) and position parameters. Notably, each particle adjusts its flight speed dynamically based on its flight history and that of its neighbors. Similarly, it attempts to adjust its location based on its current position, velocity, the distance between its current position and personal optimal, and the distance between its current position and swarm optimum.

The optimization problem is resolved when the swarm of particles (birds) arrives at the global optimum, which is the optimal solution. Following this, the mathematical models utilized to construct the PSO method, as demonstrated by the parameters and equations [39].

$$V_{i(t+1)} = WY_{i(t)} + C_1R_1(P_{best\_i} - X_{i(t)}) + C_2R_2\,(G_{best\_i} - X_{i(t)}) \tag{7}$$

In (7) is the formula for the velocity of the particle, denoted by $V_{i(t+1)}$. W is the inertia factor that controls the degree to which the particle maintains its previous velocity. The velocity of the particle i during iteration t is denoted by the variable $Y_{i(t)}$. P_{best} and G_{best}'s impact on the particle may be regulated by cognitive and social characteristics known as C_1 and C_2, respectively. The integers R_1 and R_2 are arbitrary and range from 0,1 to 1. P_{best_i} represents the highest possible position that particle i can achieve. Iteration t's location of particle i is denoted by the variable $X_{i(t)}$. G_{best} is the best position that any particle has ever managed to acquire.

The next step is to calculate the position. Each particle calculates a new position using (8). $X_{i(t+1)}$ is the position of particle i in iteration t+1. $X_{i(t)}$ is the position of particle i in iteration t. $V_{i(t+1)}$ is the velocity of particle i in iteration t+1.

$$(X_{i(t+1)} = X_{i(t)} + V_{i(t+1)} \tag{8}$$

Each particle remembers its personal best position, denoted by the notation P_{best}. Because of the P_{best} update, every particle tries to remember information on the best solution it has ever discovered; this implies that it maintains a record of the various individually effective solutions. In the meantime, G_{best} stands for the highest position that any particle in the population has ever managed to reach. Because of the G_{best} upgrade, the population's particles are now able to communicate with one another and exchange knowledge on how to achieve better results. G_{best} is a global guidance system that supplies all particles with information on the most optimum solution the population has ever discovered.

2.5.3. Grey wolf optimization

The process of parameter optimization, which involves the computation of fitness values and the adjustment of wolf locations, is illustrated in Figure 5. This figure presents the beginning stages of the GWO method. The method under consideration is a metaheuristic algorithm that draws inspiration from the hunting techniques employed by grey wolves. The three primary stages of grouping, pursuit, and engagement (GWO) emulate the inherent predatory conduct observed in a collective of grey wolves [32].

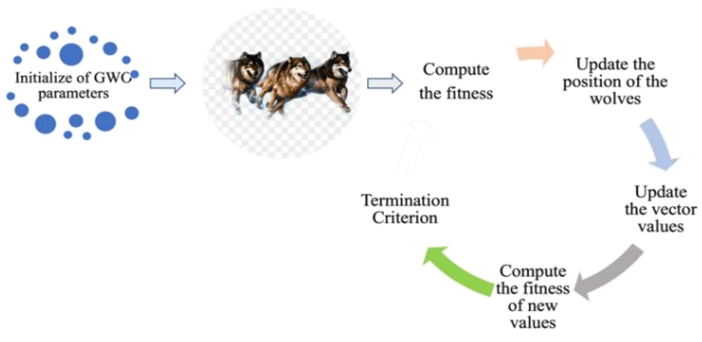


Figure 5. The GWO workflow

The leader of a pack of grey wolves is called the alpha wolf. Grey wolves live in packs. The other wolves in the pack will follow the alpha wolf's lead α. However, the beta wolves β will assist the alpha wolves in making decisions. At the third level, delta wolves δ are a safeguard to defend the group and warns the pack in the event of potential danger. Omegas ω are the very last wolves to follow the orders of senior wolves. Therefore alpha, beta, and delta are the three best solutions, and the rest of the wolves modify their positions according to rank [40]. Encircling is the process of bringing weakened wolves closer to stronger wolves to encourage exploration and the search for solutions. Here is the equation:

$$\mathrm{X}\,(\mathrm{t}+1) \;=\; \mathrm{Xpos(t)} - \mathrm{A.D} \tag{9}$$

$$A = 2ar_1 - \; a \tag{10}$$

$$C \;= 2r_2 \tag{11}$$

$$D \;=\; \left|CX_{pos}(t) \;-\; X(t)\right| \tag{12}$$

In (9)-(12), t denotes the current iteration, Ã and Ç denote coefficient vectors, X_{pos} denotes the prey's position vector, and X is the position vector of a grey wolf. The vector a value decreases linearly with the number of iterations, with t denoting the current iteration and r_1 and r_2 values being created randomly between 0 and 1. The aforementioned procedure is a fundamental component within the framework of the GWO algorithm, facilitating iterative modifications in the pursuit of optimal solutions.

In the hunting stage, this action follows and tracks the alpha, beta, and delta's wolves, leading the wolves on the prowl (possible solution). The wolves position update is described in (13):

$$X_{i(t+1)} = \; X_{i(t)} - A.D_i \tag{13}$$

Where $X_{i(t+1)}$ denotes where Wolf ì was in iteration t+1 following the hunting stage, the position of wolf ì in iteration t before the hunting stage step is represented by $X_{i(t)}$. A is the change vector resulting from the influence of alpha, beta, delta, and random factors. D_i is the Euclidean distance between the wolf position and the alpha, beta, or delta position, depending on the wolf's role.

Before an attack, keeping the space between the wolf and its prey as small as possible is important. Attacks on prey happen when the herd is upon the prey. The optimal overall solution to the optimization issue is the prey. To make a declining coefficient depending on the iteration number, it must be specified as a vector A, as illustrated in (10). As the value index α value is decreased in accordance with (14), the vector A value also decreases.

$$\alpha = 2 - \mathrm{t}\left(\frac{2}{\mathrm{T}}\right) \tag{14}$$

2.6. Testing and evaluation

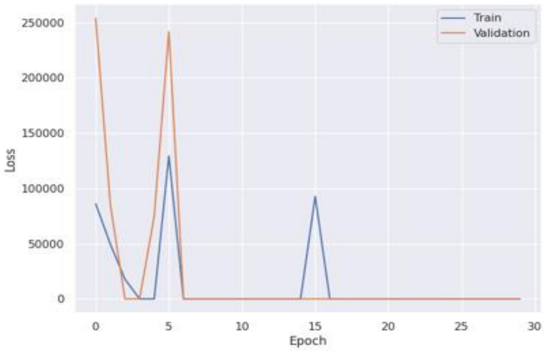
Within the phase of model testing, the test data set is employed as an assessment corpus to gauge the effectiveness of the model that has been trained beforehand. The purpose of this evaluation is to assess the model's ability to generate precise predictions for previously unseen data. The evaluation of model performance involves the utilization of several established general metrics. These metrics include mean absolute error (MAE), which quantifies the average absolute discrepancy between predicted and actual values. Additionally, root mean squared error (RMSE) provides insight into the degree of correlation between model predictions and actual data by considering the squared error. Mean absolute percentage error (MAPE) evaluates the relative error in percentage terms. Lastly, R-squared (R^2) measures the model's ability to capture variations in the data by comparing the variability of predicted results to that of the actual data. The utilization of these metrics holds significance in offering a comprehensive comprehension of the accuracy of model estimations and aiding in the evaluation of model appropriateness inside real scenarios.

3. RESULTS AND DISCUSSION

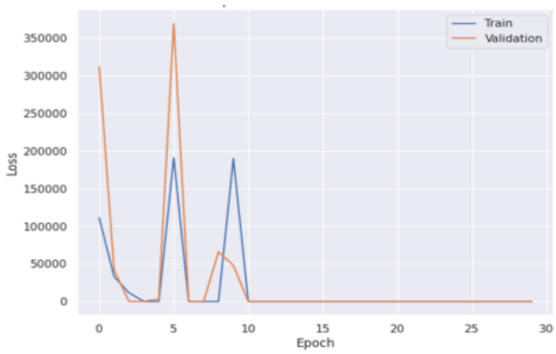
3.1. Training, validation, and testing

Randomization techniques are employed to initialize parameters within the standard LSTM model to establish a robust foundation for analysis. Figures 6(a)-(c) presents the outcomes of multiple epochs and losses during the LSTM training and validation phases were depicted separately, providing essential insights into the baseline performance of the model before undergoing optimization through the GWO and PSO algorithms. Upon optimization of the LSTM-PSO-GWO model, significant alterations were observed in the training and validation phases, as illustrated in Figures 7(a)-(c). These alterations indicate the model's enhanced capacity to converge more efficiently, as evidenced by diminishing loss value throughout multiple training iterations.

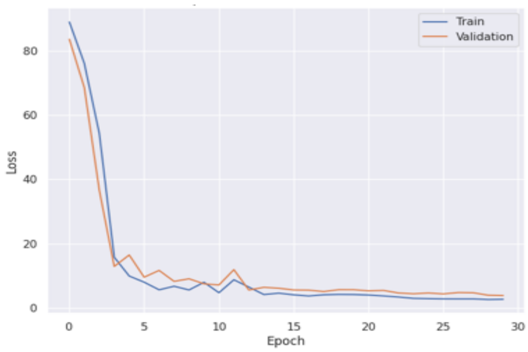
Subsequently, the evaluation progresses to the testing phases, as shown in Figures 8(a)-(c). Here, the constructed model demonstrates reduced epoch values, expediting the attainment of optimal outcomes. The decrease in loss values depicted in the graphical representation corresponds to a diminished degree of predictive discrepancy, indicating progressive convergence towards accurate predictions.



(a)

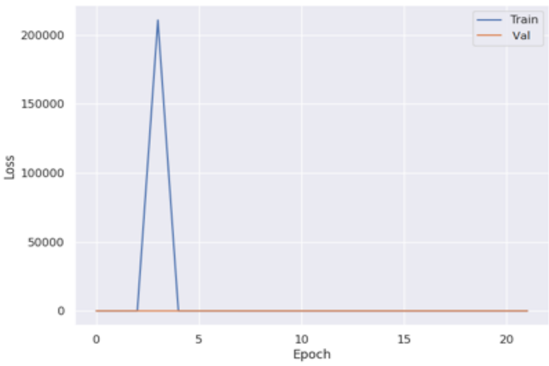


(b)

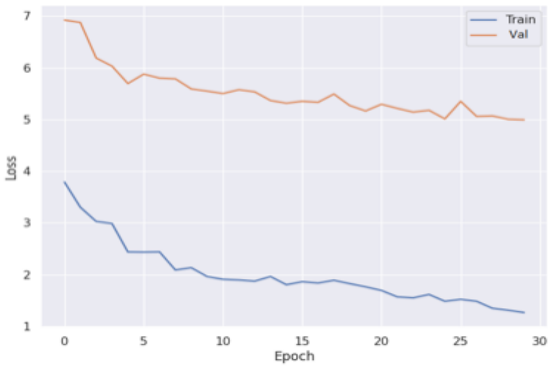


(c)

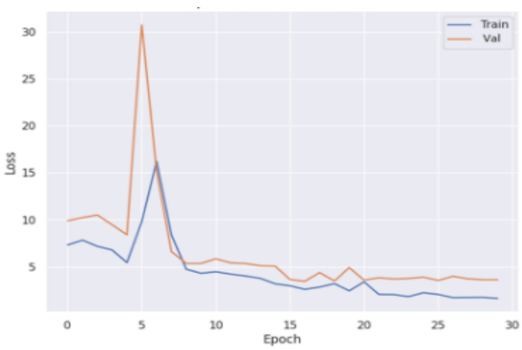
Figure 6. LSTM epoch vs loss training and validation for (a) recovery cases, (b) death cases, and (c) positive cases



(a)

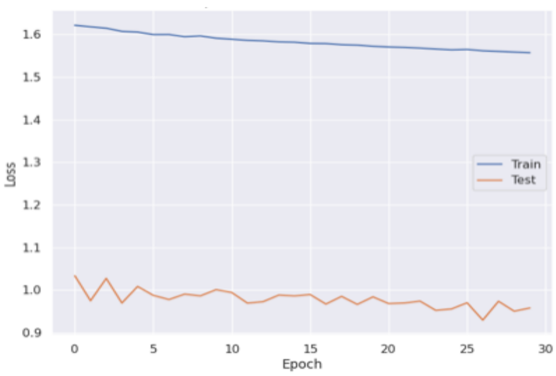


(b)

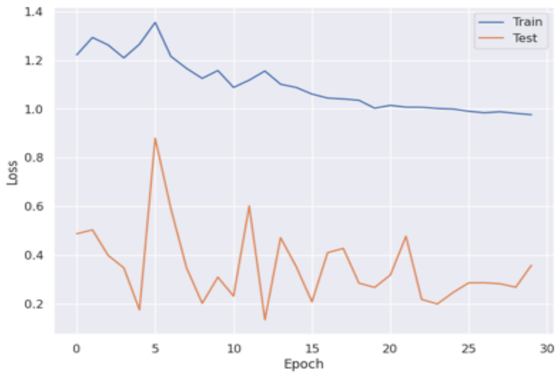


(c)

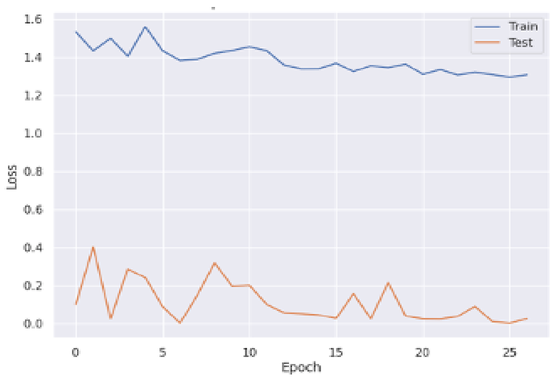
Figure 7. LSTM-PSO-GWO epoch vs loss training and validation for (a) recovery cases, (b) death cases, and (c) positive cases



(a)



(b)



(c)

Figure 8. LSTM-PSO-GWO epoch vs loss testing for (a) recovery cases, (b) death cases, and (c) positive cases

Based on the observed results depicted in the line plot, it is evident that the model's loss value exhibits a consistent decrease with each epoch. This observation suggests that the model exhibits strong performance and possesses the ability to provide accurate predictions. The findings derived from the analysis of Table 2 provide compelling evidence to support the notion that the hybrid deep learning optimization methodology, employing the LSTM-PSO-GWO model, has effectively enhanced the model's performance by diminishing loss values.

Table 2. Loss values comparison

Model	LSTM loss values	LSTM-PSO-GWO loss values
Prediction model for recovery cases	7.0	0.3
Prediction model for death cases	6.0	1.4
Prediction model for positive cases	4.0	0.1

During the initial weight LSTM stage, the utilization of the PSO method is instrumental in the pursuit of identifying an initial weight configuration that is more optimal for the LSTM model. The convergence process of the PSO algorithm assumes a crucial role in the optimization of weights by iteratively searching for a value that approximates the optimal solution within the defined search space. The convergence outcomes derived from PSO exhibit a notable trend wherein the model progressively approximates a superior solution throughout the iterative optimization procedure.

Subsequently, the GWO methodology was employed to enhance the outcomes of the optimization process executed by PSO. The GWO algorithm plays a pivotal role in the exploration of the search space, facilitating the discovery of novel solutions and enhancing the overall performance of the model. The synergistic incorporation of GWO with LSTM and PSO facilitates an enhanced capacity for the model to delve into untapped possibilities and achieve noteworthy advancements in its aptitude to discern intricate patterns from both training and validation datasets.

The analysis of the model testing outcomes demonstrates a noteworthy decrease in the MAE and RMSE metrics in comparison to the conventional LSTM model. This improvement is observed following the implementation of the suggested optimization approach for the recovery scenario, as outlined in Table 3. The decrease observed in the metrics indicates a significant enhancement in the effectiveness and accuracy of the model being evaluated. The obtained results provide confirmation that the used optimization methodology has effectively generated a superior model for tackling intricate recovery obstacles.

Table 3. Testing output prediction model for COVID-19 recovery cases

Model	MAE	RMSE	MAPE	R^2
LSTM	16788.6	18342.7	0.4	1.0
LSTM-PSO-GWO	10584.4	11324.6	0.3	1.0

This observation suggests that through the process of optimization utilizing the proposed model, the predictions generated by the model exhibit a higher degree of accuracy and proximity to the ground truth value, as evidenced by the reduction in MAE and RMSE. The empirical findings reveal that the vanilla LSTM model exhibits a MAPE of 0.4. However, upon subjecting the LSTM-PSO-GWO model to optimization, a noteworthy reduction in the MAPE value to 0.3 is observed. The observed reduction in value suggests that the LSTM-PSO-GWO model exhibits a comparatively diminished relative error rate in contrast to the vanilla LSTM model.

The utilization of the hybrid depth optimizer, in conjunction with the LSTM-PSO-GWO framework, represents a novel and cutting-edge methodology aimed at enhancing performance metrics and achieving higher levels of relative accuracy in COVID-19 death case prediction, in Table 4. By harnessing the unique capabilities and inherent benefits of the LSTM, PSO, and GWO algorithms. This novel methodology demonstrates notable enhancements in prognosticating mortality rates with heightened precision and improved efficacy.

Table 4. Testing output prediction model for COVID-19 death cases

Model	MAE	RMSE	MAPE	R^2
LSTM	3542.6	3726.4	2.5	0.5
LSTM-PSO-GWO	2047.7	2154.8	1.4	0.8

Synergistic integration of these three algorithms improves the model's ability to understand and capture complex patterns, trends, and nuances from associated data. Due to increased awareness and understanding of data variability, the model gains knowledge. This research advances awareness of mortality's complex mechanisms and multifaceted influences. The proposed model's performance, as measured by MAE, RMSE, MAPE, and R^2, improves significantly after implementation. The combination of these three techniques reduces model errors and improves mortality prediction accuracy.

The findings presented in Table 5 demonstrate the outcomes obtained from evaluating the LSTM-PSO-GWO hybrid deep optimization model on positive predictions. Notably, the accuracy of the model exhibits improvement, as evidenced by a reduction in the MAE and RMSE values. The findings of this study suggest that the optimized model exhibits enhanced predictive capabilities, yielding more accurate and proximate estimations for positive predictions.

Table 5. Testing output prediction model for COVID-19 positive cases

Model	MAE	RMSE	MAPE	R^2
LSTM	30.5	30.5	0.1	0.0
LSTM-PSO-GWO	28.3	28.3	0.1	0.0

3.2. Forecasting

The hybrid model that has been developed to forecast the number of cases recovered from COVID-19 for the upcoming seven days has produced its output, in Table 6. The aforementioned observation implies that the hybrid forecasting model demonstrates a propensity for projecting a surge in the count of individuals recuperating from the COVID-19 ailment within the specified temporal interval. In addition to the prevailing trend, there appears to be a consistent and enduring upward trajectory. The observed trend in the daily increments of recovered cases exhibited a notable degree of consistency.

The hybrid forecasting model demonstrates the ability to predict an increase in COVID-19 pandemic fatalities throughout seven days, as indicated by the analysis in Table 7. The increase in mortality rates can be attributed to several factors, such as the proliferation of highly contagious virus strains, the vulnerability of specific demographic cohorts, and potential delays in providing adequate treatment for more critical patients. Therefore, it is imperative to be vigilant and implement appropriate measures to effectively manage these sudden increases.

Table 6. Forecast the number of recovery cases

	2020-04-01	2020-04-02	2020-04-03	2020-04-04	2020-04-05	2020-04-06	2020-04-07
0	3414109	3455376	3484840	3519976	3556195	3594315	3629528
1	3443903	3481798	3510439	3545982	3579195	3615706	3648946
2	3472915	3508083	3538597	3571287	3603999	3638242	3671504
3	3499037	3536199	3563767	3595947	3626320	3660606	3691176
4	3522048	3560014	3584181	3613724	3641954	3674126	3701492
...	...	...	...	...	...	...	...
119	4110574	4097943	4094781	4084359	4073234	4056569	4041404
120	4110811	4098194	4095056	4084689	4073529	4056907	4041731
121	4111045	4098464	4095372	4084967	4073843	4057212	4042080
122	4111250	4098762	4095596	4085223	4074064	4057483	4042308
123	4111464	4098964	4095783	4085368	4074223	4057609	4042420

Table 7. Forecast the number of died cases

	2020-04-01	2020-04-02	2020-04-03	2020-04-04	2020-04-05	2020-04-06	2020-04-07
0	120013	120527	121189	121914	122903	123969	125210
1	121141	121739	122480	123252	124253	125324	126541
2	122633	123280	124075	124760	125739	126707	127892
3	123981	124702	125405	126136	127032	128023	129236
4	125342	125978	126725	127394	128319	129355	130623
...	...	...	...	...	...	...	...
119	143960	146468	148716	151025	152450	153615	154088
120	143969	146476	148723	151031	152456	153623	154097
121	143979	146485	148731	151039	152465	153633	154108
122	143986	146492	148737	151047	152474	153643	154118
123	143998	146502	148750	151059	152487	153654	154129

A discernible and persistent trend emerges in Table 8, whereby the quantity of positive cases exhibits a steady and incremental rise from the initial day to the fourteenth day. The observed trend in the incremental increase of positive cases over consecutive days suggests a consistent and stable pattern, thereby implying a sustained progression in the number of cases. While the hybrid model has demonstrated commendable consistency in its prognostications regarding the escalation of positive cases, it is important to acknowledge that the forecasted outcomes may exhibit marginal disparities when juxtaposed with the forthcoming actual data. The observed disparity can be attributed to a multitude of variables, encompassing alterations in pandemic mitigation strategies, fluctuating rates of immunization, or unanticipated elements that exert an influence on the transmission dynamics of the virus. Consequently, it is imperative to maintain a regular cadence of updates to the forecasts to assimilate the most recent data and mitigate the occurrence of prediction errors.

Table 8. Forecast the number of positive cases

	2020-04-01	2020-04-02	2020-04-03	2020-04-04	2020-04-05	2020-04-06	2020-04-07
0	51318	51346	51361	51368	51377	51389	51395
1	51318	51346	51361	51368	51377	51389	51395
2	51318	51346	51361	51368	51377	51389	51395
3	51318	51346	51361	51368	51377	51389	51395
4	51318	51346	51361	51368	51377	51389	51395
...	...	...	...	...	...	...	...
119	51318	51346	51361	51368	51377	51389	51395
120	51318	51346	51361	51368	51377	51389	51395
121	51318	51346	51361	51368	51377	51389	51395
122	51318	51346	51361	51368	51377	51389	51395
123	51318	51346	51361	51368	51377	51389	51395

4. CONCLUSION

Based on the empirical evidence, the combination of PSO and GWO techniques is employed in tandem to effectively optimize the parameters of the LSTM model. The hybrid deep learning optimization model, which combines LSTM, PSO, and GWO techniques, demonstrates superior performance in terms of reduced loss values, accelerated convergence, and absence of overfitting. This approach holds the potential to enhance the precision of COVID-19 recovery prediction, mortality prediction, and positive case prediction. However, a re-evaluation is necessary, particularly in terms of incorporating a more informative feature selection algorithm to account for differences in objective data. The model under consideration exhibits significant potential for applicability across diverse sectors and can function as a beneficial instrument for facilitating intelligent and efficient decision-making procedures.

ACKNOWLEDGEMENTS

The authors would like to express gratitude to MIIT, Universiti of Kuala Lumpur Malaysia, and Universitas Duta Bangsa Surakarta for their invaluable support and financial aid in facilitating the research and publication endeavors. In particular, thanks to Universitas Duta Bangsa Surakarta for grant number 049/UDB.LPPM/A.34-HK/II/2023, as well as to Universiti of Kuala Lumpur Malaysia for funding the publication of this article. The provision of payment helps and facilities has greatly contributed to the advancement of research and publication endeavors.

REFERENCES

[1] A. Zumla and D. S. C. Hui, "Emerging and reemerging infectious diseases: global overview," *Infectious Disease Clinics of North America*, vol. 33, no. 4, pp. 13–19, 2019, doi: 10.1016/j.idc.2019.09.001.
[2] P. Chen *et al.*, "A study of the relationship between human infection with avian influenza a (H5N6) and environmental avian influenza viruses in Fujian, China," *BMC Infectious Diseases*, vol. 19, no. 1, pp. 1-8, 2019, doi: 10.1186/s12879-019-4145-6.
[3] X. Wu, L. Xiao, and L. Li, "Research progress on human infection with avian influenza H7N9," *Frontiers of Medicine*, vol. 14, no. 1, pp. 8–20, 2020, doi: 10.1007/s11684-020-0739-z.
[4] A. K. Khetan, "COVID-19: Why declining biodiversity puts us at greater risk for emerging infectious diseases, and what we can do," *Journal of General Internal Medicine*, vol. 35, no. 9, pp. 2746–2747, 2020, doi: 10.1007/s11606-020-05977-x.
[5] D. Chang *et al.*, "Epidemiologic and clinical characteristics of novel coronavirus infections involving 13 patients outside Wuhan, China," *JAMA - Journal of the American Medical Association*, vol. 323, no. 11, pp. 1092–1093, 2020, doi: 10.1001/jama.2020.1623.
[6] S. Dana, A. B. Simas, B. A. Filardi, R. N. Rodriguez, L. L. da C. Valiengo, and J. Gallucci-Neto, "Brazilian modeling of COVID-19 (BRAM-COD): a Bayesian Monte Carlo approach for COVID-19 spread in a limited data set context," *medRxiv*, vol. 19, pp. 1-42, doi: 10.1101/2020.04.29.20081174.
[7] S. Mirri, G. Delnevo, and M. Roccetti, "Is a COVID-19 second wave possible in Emilia-Romagna (Italy)? Forecasting a future outbreak with particulate pollution and machine learning," *Computation*, vol. 8, no. 3, pp. 1-25, 2020, doi: 10.3390/computation8030074.
[8] M. Schuit *et al.*, "Airborne SARS-CoV-2 is rapidly inactivated by simulated sunlight," *Journal of Infectious Diseases*, vol. 222, no. 4, pp. 564–571, 2020, doi: 10.1093/infdis/jiaa334.
[9] S. Ibrahim, "The influences of global geographical climate towards COVID-19 spread and death," *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 9, no. 1.4, pp. 612–617, 2020, doi: 10.30534/ijatcse/2020/8591.42020.
[10] I. Al-Turaiki, M. Alshahrani, and T. Almutairi, "Building predictive models for MERS-CoV infections using data mining techniques," *Journal of Infection and Public Health*, vol. 9, no. 6, pp. 744–748, 2016, doi: 10.1016/j.jiph.2016.09.007.
[11] N. L. Bragazzi, H. Dai, G. Damiani, M. Behzadifar, M. Martini, and J. Wu, "How big data and artificial intelligence can help better manage the covid-19 pandemic," *International Journal of Environmental Research and Public Health*, vol. 17, no. 9, pp. 1-8, 2020, doi: 10.3390/ijerph17093176.
[12] F. Petropoulos and S. Makridakis, "Forecasting the novel coronavirus COVID-19," *PLoS ONE*, vol. 15, no. 3, pp. 1-8, 2020, doi: 10.1371/journal.pone.0231236.
[13] Z. Yang *et al.*, "Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health interventions," *Journal of Thoracic Disease*, vol. 12, no. 3, pp. 165–174, 2020, doi: 10.21037/jtd.2020.02.64.
[14] J. Milojković, M. Milić, and V. Litovski, "ANN model for one day ahead Covid-19 prediction," *International Conference IcETRAN*, pp. 1–4, 2022.
[15] A. N. A. Kamarudin, Z. Zainol, N. F. A. Kassim, and R. Sharif, "Prediction of COVID-19 cases in Malaysia by using machine learning: A preliminary testing," *2021 International Conference of Women in Data Science at Taif University, WiDSTaif 2021*, 2021, pp. 1-6, doi: 10.1109/WIDSTAIF52235.2021.9430222.
[16] G. K and V. B, "An efficient COVID-19 pandemic survival analysis to compare random forest and support vector machine for classifying performance in censored data," *ECS Transactions*, vol. 107, no. 1, pp. 12993–13004, 2022, doi: 10.1149/10701.12993ecst.
[17] S. Jain, A. Dhall, S. Patiyal, and G. P. S. Raghava, "IL13Pred: A method for predicting immunoregulatory cytokine IL-13 inducing peptides," *Computers in Biology and Medicine*, vol. 143, pp. 1-8, Apr. 2022, doi: 10.1016/j.compbiomed.2022.105297.
[18] W. A. Wali, "Application of particle swarm optimization with ANFIS model for double scroll chaotic system," *International Journal of Electrical and Computer Engineering*, vol. 11, no. 1, pp. 328–335, 2021, doi: 10.11591/ijece.v11i1.pp328-335.
[19] N. B. Yahia, M. D. Kandara, and N. B. B. Saoud, "Deep ensemble learning method to forecast COVID-19 outbreak," *Research Square*, pp. 1–20, 2020, doi: 10.21203/rs.3.rs-27216/v1.
[20] S. F. Ardabili *et al.*, "COVID-19 outbreak prediction with machine learning," *Algorithms*, vol. 13, no. 10, pp. 1-36, 2020, doi: 10.3390/a13100249.
[21] L. J. Muhammad, E. A. Algehyne, S. S. Usman, A. Ahmad, C. Chakraborty, and I. A. Mohammed, "Supervised machine learning models for prediction of COVID-19 infection using epidemiology dataset," *SN Computer Science*, vol. 2, no. 1, pp. 1-13, 2021, doi: 10.1007/s42979-020-00394-7.
[22] 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 and Fractals*, vol. 139, pp. 1-9, 2020, doi: 10.1016/j.chaos.2020.110017.
[23] V. K. R. Chimmula and L. Zhang, "Time series forecasting of COVID-19 transmission in Canada using LSTM networks," *Chaos, Solitons and Fractals*, vol. 135, pp. 1-6, 2020, doi: 10.1016/j.chaos.2020.109864.
[24] M. Masum, H. Shahriar, H. M. Haddad, and M. S. Alam, "R-LSTM: time series forecasting for COVID-19 confirmed cases with LSTMbased framework," *Proceedings - 2020 IEEE International Conference on Big Data, Big Data 2020*, pp. 1374–1379, 2020, doi: 10.1109/BigData50022.2020.9378276.
[25] N. Ayoobi *et al.*, "Time series forecasting of new cases and new deaths rate for COVID-19 using deep learning methods," *Results in Physics*, vol. 27, pp. 1-15, 2021, doi: 10.1016/j.rinp.2021.104495.
[26] J. Devaraj *et al.*, "Forecasting of COVID-19 cases using deep learning models: Is it reliable and practically significant?," *Results in Physics*, vol. 21, pp. 1-25, 2021, doi: 10.1016/j.rinp.2021.103817.
[27] M. O. Alassafi, M. Jarrah, and R. Alotaibi, "Time series predicting of COVID-19 based on deep learning," *Neurocomputing*, vol. 468, pp. 335–344, 2022, doi: 10.1016/j.neucom.2021.10.035.
[28] H. Gupta *et al.*, "Data analytics and mathematical modeling for simulating the dynamics of COVID-19 epidemic—a case study of India," *Electronics*, vol. 10, no. 2, pp. 1–21, 2021, doi: 10.3390/electronics10020127.
[29] A. Kaveh, "Particle swarm optimization," *Advances in Metaheuristic Algorithms for Optimal Design of Structures*, pp. 11–43,

2017, doi: 10.1007/978-3-319-46173-1_2.

[30] Y. Zhang and S. Yang, "Prediction on the highest price of the stock based on PSO-LSTM neural network," *2019 IEEE 3rd International Conference on Electronic Information Technology and Computer Engineering, EITCE 2019*, pp. 1565–1569, 2019, doi: 10.1109/EITCE47263.2019.9094982.

[31] W. Lu, W. Jiang, N. Zhang, and F. Xue, "Application of PSO-based LSTM neural network for outpatient volume prediction," *Journal of Healthcare Engineering*, vol. 2021, pp. 1-10, 2021, doi: 10.1155/2021/7246561.

[32] J. Aguila-Leon, C. Chiñas-Palacios, C. Vargas-Salgado, E. Hurtado-Perez, and E. X. M. Garcia, "Particle swarm optimization, genetic Algorithm and grey Wolf optimizer algorithms performance comparative for a DC-DC boost converter PID controller," *Advances in Science, Technology and Engineering Systems*, vol. 6, no. 1, pp. 619–625, 2021, doi: 10.25046/aj060167.

[33] I. M. El-Hasnony, S. I. Barakat, and R. R. Mostafa, "Optimized ANFIS model using hybrid metaheuristic algorithms for parkinson's disease prediction in IoT environment," *IEEE Access*, vol. 8, pp. 119252–119270, 2020, doi: 10.1109/ACCESS.2020.3005614.

[34] S. Yan, "Understanding LSTM and its diagrams," *Medium*, pp. 1–12, 2016, [Online]. Available: https://medium.com/mlreview/understanding-lstm-and-its-diagrams-37e2f46f1714.

[35] X. H. Le, H. V. Ho, G. Lee, and S. Jung, "Application of long short-term memory (LSTM) neural network for flood forecasting," *Water*, vol. 11, no. 7, pp. 1-19, 2019, doi: 10.3390/w11071387.

[36] H. R. R. Zaman and F. S. Gharehchopogh, "An improved particle swarm optimization with backtracking search optimization algorithm for solving continuous optimization problems," *Engineering with Computers*, vol. 38, pp. 2797–2831, 2022, doi: 10.1007/s00366-021-01431-6.

[37] Y. Wu and Q. Song, "Improved particle swarm optimization algorithm in power system network reconfiguration," *Mathematical Problems in Engineering*, vol. 2021, pp. 1-11, 2021, doi: 10.1155/2021/5574501.

[38] C. Y. Tsai and I. W. Kao, "Particle swarm optimization with selective particle regeneration for data clustering," *Expert Systems with Applications*, vol. 38, no. 6, pp. 6565–6576, 2011, doi: 10.1016/j.eswa.2010.11.082.

[39] J. K. and R. Eberhart, "Particle swarm optimization," *Proceedings of ICNN'95 - International Conference on Neural Networks*, vol. 4, pp. 1942–1948, 1995, doi: 10.1109/ICNN.1995.488968.

[40] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey wolf optimizer," *Advances in Engineering Software*, vol. 69, pp. 46–61, 2014, doi: 10.1016/j.advengsoft.2013.12.007.

BIOGRAPHIES OF AUTHORS

Faulinda Ely Nastiti received the master's degree with a specialization in Information Technology, from the Faculty of Engineering Universitas Gadjah Mada, Indonesia in 2012. She is currently pursuing the Ph.D. degree in information technology with Universiti Kuala Lumpur with a focus on deep learning. She was a certified data center professional and big data certified associate. She is a senior lecturer and researcher at Universitas Duta Bangsa Surakarta, Indonesia. She is authored more than 40 research papers published in journals and conferences. Her areas of interest include machine learning, deep learning, and intelligence systems. She can be contacted at email: nastiti.faulinda@s.unikl.edu.my or faulinda_ely@udb.ac.id.

Prof. Dr. Shahrulniza Musa received the master's degree in integrated research study and the Ph.D. degree in Communication Network Security from the Faculty of Electrical and Electronic Engineering, Loughborough University, U.K., in 2005 and 2008, respectively. He is a full professor at the Malaysian Institute of Information Technology (MIIT), Universiti Kuala Lumpur (UniKL). Apart from teaching and post-graduate supervision, he is also active in software project consultation and development in business applications, enterprise resource planning (ERP), and customer relation management (CRM). His research interests are in cybersecurity, IoT security, big data analytics, and SDN. He can be contacted at email: shahrulniza@unikl.edu.my.

Eiad Yafi, P.hD. received the master's degree in integrated research study and the Ph.D. degree in Computer Science from the Jamia Hamdard University, New Delhi India, in 2004 and 2011. He is visiting professor at the University of Technology Sydney, Australia. His research interest includes data science, artificial intelligence, knowledge discovery, and information retrieval: social media analysis, and process mining in health and financial institution. He is senior member of IEEE. He can be contacted at email: eiad.yafi@gmail.com.