IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 5 January 2024, accepted 18 February 2024, date of publication 21 February 2024, date of current version 28 February 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3368034

RESEARCH ARTICLE

Fine-Tuning of Predictive Models CNN-LSTM and CONV-LSTM for Nowcasting $PM_{2.5}$ Level

TAFIA HASNA PUTRI[1], REZZY EKO CARAKA[1,2,3,4], (Member, IEEE), TONI TOHARUDIN[1], YUNHO KIM[3], RUNG-CHING CHEN[5], (Member, IEEE), PRANA UGIANA GIO[6], ANJAR DIMARA SAKTI[7], RESA SEPTIANI PONTOH[1], INDAH RESKI PRATIWI[1], FARID AZHAR LUTFI NUGRAHA[1], THALITA SAFA AZZAHRA[1], JESSICA JESSLYN CERELIA[1], GUMGUM DARMAWAN[1], DEFI YUSTI FAIDAH[1], AND BENS PARDAMEAN[8,9]

[1]Department of Statistics, Faculty of Mathematics and Natural Science, Padjadjaran University, Sumedang, West Java 45361, Indonesia
[2]Research Center for Data and Information Sciences, Research Organization for Electronics and Informatics, National Research and Innovation Agency, Bandung, West Java 40135, Indonesia
[3]Department of Mathematical Sciences, Ulsan National Institute of Science and Technology, Ulsan 44919, Republic of Korea
[4]School of Economics and Business, Telkom University, Bandung, West Java 40257, Indonesia
[5]Department of Information Management, Chaoyang University of Technology, Taichung 41349, Taiwan
[6]Department of Mathematics, Universitas Sumatera Utara, Medan 20155, Indonesia
[7]Remote Sensing and Geographic Information Sciences Research Group, Faculty of Earth Sciences and Technology, Institut Teknologi Bandung, Bandung 40132, Indonesia
[8]Bioinformatics and Data Science Research Center, Bina Nusantara University, Jakarta 11480, Indonesia
[9]Computer Science Department, BINUS Graduate Program—Master of Computer Science Program, Bina Nusantara University, Jakarta 11480, Indonesia

Corresponding authors: Rezzy Eko Caraka (rezzy.eko.caraka@brin.go.id) and Yunho Kim (yunhokim@unist.ac.kr)

The work of Toni Toharudin, Resa Septiani Pontoh, and Gumgum Darmawan was supported by the Directorate of Research and Community Engagement, Universitas Padjadjaran. The work of Rezzy Eko Caraka was supported in part by the National Research Foundation of Korea under Grant NRF-2023R1A2C1006845, and in part by Telkom University. The work of Yunho Kim was supported by the National Research Foundation of Korea under Grant NRF-2022R1A5A1033624 and Grant NRF-2023R1A2C1006845. The work of Rung-Ching Chen was supported by the Ministry of Science and Technology, Taiwan, under Grant NSTC-111-2221-E-324-020, Grant NSTC-111-2622-E-324-002, Grant NSTC-112-2221-E-324-003-MY3, and Grant NSTC-112-2221-E-324-011-MY2.

ABSTRACT Particulate matter forecasting is fundamental for early warning and controlling air pollution, especially $PM_{2.5}$. The increase in this level of concentration will lead to a negative impact on public health. This study develops a hybrid model of CNN-LSTM and CONV-LSTM by combining a convolutional neural network (CNN) with an LSTM network to forecast $PM_{2.5}$ concentration for the next few hours in Kemayoran DKI Jakarta, which is known as a busy area. We discovered the advantages of CNN in effectively extracting features and LSTM in learning long-term historical data from $PM_{2.5}$ concentration time series data. The predictive model of CNN-LSTM is carried out in a different architecture where the CNN process is carried out first to become the input of LSTM. For CONV-LSTM, it is carried out in one architecture where the multiplication in the LSTM architecture is coupled with the convolution process. This research will explain how the method of developing hybrid CNN-LSTM and CONV-LSTM in predicting $PM_{2.5}$ concentrations. Based on metric evaluation, the two models are compared to find the best model. Both predictive models produce MAPE values that fall into the good enough category with values <20%. Results were obtained for CONV-LSTM with MAE worth 6.52, RMSE 8.55, and MAPE 16.39%. As a result, the CONV-LSTM model performs better than CNN-LSTM in nowcasting $PM_{2.5}$.

INDEX TERMS $PM_{2.5}$, time series, CNN, LSTM, nowcasting.

I. INTRODUCTION

The World Health Organisation (WHO) points out that air pollution has overwhelmed human life from various directions [1], [2], [3], [4]. Air pollution is a severe problem

The associate editor coordinating the review of this manuscript and approving it for publication was Mingbo Zhao.

in big cities in Indonesia. There are six main types of pollutants based on the World Air Quality Index (WAQI), which include Ozone (O3), Nitrogen Dioxide (NO_2), Sulphur Dioxide (SO_2), Carbon Monoxide (CO), $PM_{2.5}$, and PM_{10} emissions. The main concern, however, is the content of particulate Matter (PM) 2.5 as it is one of the most dangerous types of major pollutants if it exceeds the safe limit of the World Health Organization standard when the concentration is less than 25 μg/m^3 [5], [6].

$PM_{2.5}$ is a very small air pollutant, about 2.5 micrometers or less in diameter, which is smaller than 3% of the diameter of a human hair. PM, also known as particle pollution, constitutes a blend of solid and liquid particles present in the air. This amalgamation encompasses particles such as dust, dirt, soot, and smoke. Prolonged exposure to heightened levels of PM2.5 is linked to a spectrum of respiratory issues, including exacerbated asthma, bronchitis, and other respiratory as well as cardiovascular diseases [2], [3], [4], [6], [7].

Jakarta is the national capital of Indonesia, well-known as the nation's economic, political, and cultural center, with a metropolitan area of 6392 m^2 [8], [9], [10] It is reported that Jakarta's air quality is inferior, with many factors contributing to the high pollution in Jakarta [3]. The Meteorology, Climatology, and Geophysics Agency (BMKG) has recorded that the decline in air quality in the Jakarta area is caused by conducive meteorological factors that cause the accumulation of $PM_{2.5}$ concentrations. The Kemayoran area in Jakarta shows that throughout June 2022, the average concentration of $PM_{2.5}$ was 41 μg/m^3, which is included in the moderate category. Specifically, the Kemayoran area contributed the highest pollution with 169 US AQI, equal to 90 μg/ m^3, followed by Pejaten Barat with 155 US AQI or 63.2 μg/ m^3 [11]. In third place, the US Embassy in Central Jakarta touched 153 US AQI or 59.3 μg/ m^3.

Therefore, we investigate $PM_{2.5}$ forecasting in the Kemayoran area with the next hour's output. The prediction results can help prevent public health from the adverse effects of air pollution. Apart from the people, this real-time prediction allows more rapid decision-making in many sectors, such as transport, energy, and industry [12], [13], [14]. Using real-time $PM_{2.5}$ prediction, companies can reduce production or postpone activities that produce air pollutant emissions. The best time to predict $PM_{2.5}$ is within the next 24 hours since the further ahead the prediction is; the more likely weather and pollution patterns will affect it. By predicting $PM_{2.5}$ for the next approximately 24 hours, the timeframe is sufficient to provide information for the public to take preventive actions like avoiding outdoor activities and using air masks if $PM_{2.5}$ concentrations are expected to be elevated [15], [16], [17], [18].

Predictions on a narrow domain interval (24 hours ahead) require detailed and accurate observation data. $PM_{2.5}$ concentration data has a large amount of historical data and tends to have high volatility or rapid and significant fluctuations in variable levels, which can be challenging to estimate data trends and patterns [19]. It is also known that $PM_{2.5}$ concentrations show a diurnal pattern indicating the difference between day and night, where the data tends to increase in the early mornings and decrease in the afternoons and the evenings, revealing a complex relationship between time of day and $PM_{2.5}$ concentrations [20], [21], [22], [23].

Deep learning is an advanced machine learning implementation method based on artificial neural networks, popularly adopted in the past few years. Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN) are commonly used for pattern detection [24], [25], [26], object detection [27], [28], [29], image classification [30], [31], [32], and other purposes. At the same time, RNN has shortcomings, especially the problem of long-term dependence on time series data which causes loss of gradient, leading to the formation of the Long Short-Term Memory (LSTM) algorithm, which is a development of RNN in overcoming these problems. Data growth requires more complex analysis models, such as hybrid deep learning models, CNN, and LSTM, for forecasting [33], [34], [35], [36]. CNN can perform feature extraction on the model, and LSTM works in predicting data over a long period. For this reason, this study uses CNN-LSTM and CONV-LSTM methods in predicting $PM_{2.5}$, where the results of the two approaches will be compared based on the evaluation of the specified model metrics.

This paper presents novel predictive hybrid models designed to address the challenge of forecasting hourly $PM_{2.5}$ concentrations. Our innovative approach leverages air quality observation data collected from the Kemayoran BMKG station in Central Jakarta during the period from 21 May to 21 June 2022 for model development. This study makes several noteworthy contributions. Primarily, a comparative analysis of two distinct hybrid methods, CNN-LSTM and CONV-LSTM, sheds light on their effectiveness and relative advantages in predicting $PM_{2.5}$ concentrations, thereby providing valuable insights for further advancements. Secondly, the use of LSTM in $PM_{2.5}$ prediction is shown to be advantageous in handling long-term temporal dependencies and capturing historical information within the data sequence to discern and model temporal patterns. The convolution process in CNN-LSTM and CONV-LSTM further augments the prediction accuracy. Finally, the application of the best-performing method yields precise predictions for the next 24 hours in the Kemayoran area, with minimal errors. This outcome holds significant potential for assisting stakeholders, including environmental agencies, government entities, and the general public, in implementing more effective measures to mitigate exposure to air pollution. The distinctiveness of our models lies in its ability to offer enhanced prediction accuracy through the integration of CNN-LSTM and CONV-LSTM, thereby contributing to the ongoing discourse on the $PM_{2.5}$

prediction models. The remainder of the paper is organized as follows. "Methodology" section reviews recent and popular statistics and data science methods for forecasting and nowcasting. "Discussion" section presents our dataset and research location. "Results and Discussion" describes descriptive statistics and analysis using our proposed methods. Finally, section "Practical Implication" and "Conclusion".

II. METHODOLOGY

There are three main ways to predict air quality: numerical modeling, statistical modeling, and artificial intelligence (AI) methods. In numerical modeling, it usually solves very complex differential equations which require modeling procedures with considerable time and computational cost [37]. Statistical modeling makes use of collected data under statistical assumptions and properties of the data [38], [39], [40]. One classical statistical modeling often used is Autoregressive Integrated Moving Average (ARIMA). Several assumptions such as stationarity are not satisfied in practice, which makes it pretty challenging to identify non-linear relationships in the data [41].

The other approach to predicting air quality is using AI algorithms. Machine learning algorithms such as Support Vector Regression (SVR) [42], [43], [44], Random Forest (RF) [45], [46], [47], [48], Extreme Gradient Boosting (XGBoost) [49], [50], [51], [52], and Artificial Neural Networks (ANNs) have shown their applicability to air quality prediction. Among thease, conventional statistical models such as SVR [53], [54], [55], RF Pipeline (RFP) [11], [56], [57], ARIMA [58], [59], Seasonal ARIMA(SARIMA) [60], [61], and Multi-Layer Perceptron (MLP) [62], [63], [64], [65] have lower prediction evaluation values than technology-based neural network methods [66], [67], [68], [69].

In fact, this neural network based approach can handle complex non-linear relationships in the model, robust to noise in the data. With the advancement of AI algorithms, deep neural networks (DNNs) have become a promising option for predicting air quality, which is because deeper and wider networks for complex data analysis are required for bigger data size. One can find a few works in the literature about predicting time series data using machine learning. However, [70] shows the superiority of an artificial neural network (ANN) method in predicting $PM_{2.5}$ concentrations in Delhi. However, ANN is still not good enough in dealing with time series data having repeating patterns because it does not remember previous time patterns. As a evidence, [71] used LSTM for a different task of predicting air pollution concentrations with various method comparisons. In line with this, [72], [73] predicted stock prices and air quality indices by comparing several deep learning methods. The results of these two studies confirmed the best performance of the CNN-LSTM hybrid approach for predicting stock prices and air prices quality index.

A. PRE-PROCESSING

In this research, we use essential stages, including data preprocessing, handling missing data, scaling the dataset, dividing training and testing data, modeling CNN-LSTM and CONV-LSTM, selecting the best model, and making predictions. The data preprocessing stage is to identify anomalies in the data and handle data imputation, followed by data scaling using z-score. Immediately after the data preprocessing stage, data splitting or partitioning will be carried out in three parts: training, validation, and test data. Data separation is carried out with four scenarios with different proportions, specifically 90:10, 80:20, 70:30, and 60:40 [74], [75], [76], [77].

In the next stage, predictive model modeling using CNN-LSTM and CONV-LSTM, with each step of the predictive model, the CNN-LSTM method will be processed in several CNN layers first so that the output of the CNN becomes the input for the LSTM process, while for the CONV-LSTM method is carried out in the same architecture as LSTM so that the data splitting process will be the input for the CONV-LSTM process. After modeling the predictive model, the model will be trained until the loss in the model reaches an optimal or convergent point; if it has been achieved, the next step is to evaluate the model using three evaluation metrics, namely RMSE, MAE, and MAPE. In reaching the goal in this research, an increase in the trained model is carried out with a maximum iteration value until it produces a MAPE value of <20% (the forecasting model category is quite good). The last stage in this research is to predict $PM_{2.5}$ concentrations based on the model that has the minimum metric evaluation value or the minimum metric value.

The LSTM architecture has three gates, each having a process to protect and control states which are horizontal lines with the ability to all output layers in the LSTM [85]. Forget gates that determine which information should be retained and discarded from cell states; selecting information that is retained reduces the amount of information that must be passed and processed in each layer so that forget gates can help overcome the vanishing gradient problem [80], [86], [87], [88], [89]. The input gate consists of two parts; the first part uses a sigmoid function to determine which information is updated, and the second part uses a tanh process to determine the vector to be added to the cell state. The next step is to determine the output result, where the sigmoid layer determines the part of the cell state that will be output.

$$f_t = \sigma(W_f . [h_{t-1}, x_t] + b_f) \tag{1}$$

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

$$C_t = f_t * C_{t-1} + i_t * C_t \tag{3}$$

$$o_t = \sigma\ (W_o . [h_{t-1}, x_t] + b_o) \tag{4}$$

where σ represent activation function, t represents the current time statet, $t - 1$ represents the previous time state, X represents input, H represents output, and $\boldsymbol{W_f}, \boldsymbol{W_i}, \boldsymbol{W_C}$

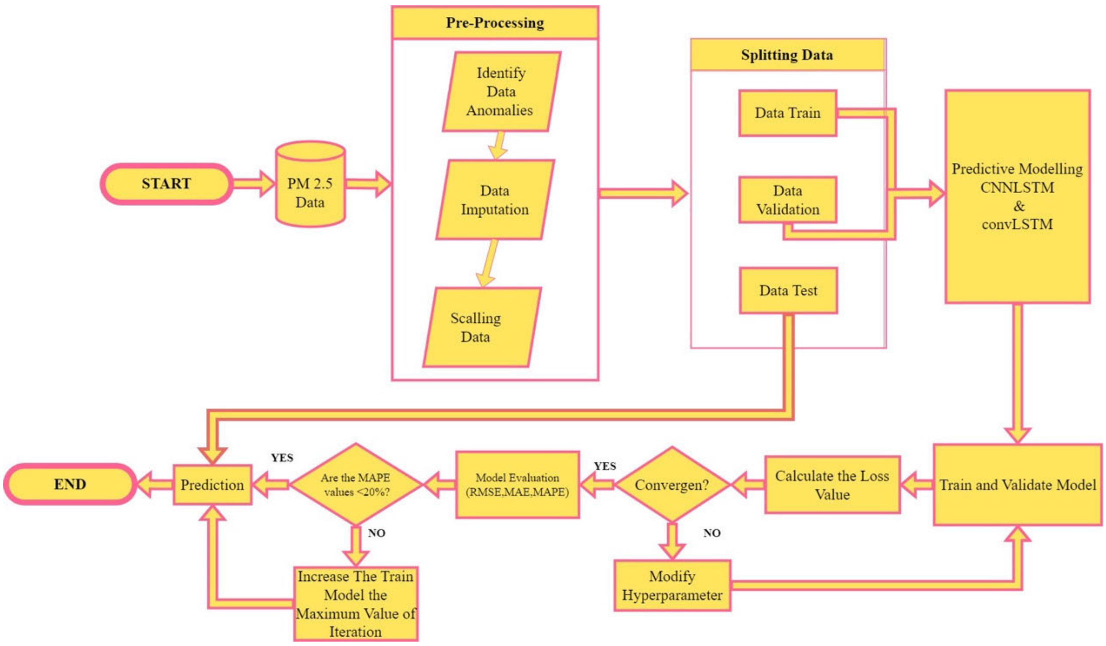


FIGURE 1. Research flowchart.

and $\boldsymbol{W_o}$ are input weights, $\boldsymbol{b_f}$, $\boldsymbol{b_i}$, $\boldsymbol{b_c}$ and $\boldsymbol{b_o}$ are bias weights. Each gate in the LSTM architecture has a weight that can be adjusted during the training process, thus helping the LSTM learn to organize the information received and stored in the memory cells. Utilizing gates, memory cells, and states, the LSTM can overcome the vanishing gradient problem of traditional RNNs and learn to remember information over a longer time [36]. As in the forget gates process, which determines which information should be retained and discarded from the cell sites by the sigmoid layer, which produces an output number between 0 and 1 to control how much information will be kept in long-term memory and how much information will be passed to the LSTM output, by selecting the retained information it reduces the amount of information that must be passed and processed at each layer so that forget gates can help overcome the vanishing gradient problem.

B. CNN FOR TIME SERIES APPLICATION

A CNN architecture is applied to $PM_{2.5}$ concentration data where n is the length of the time series and k the number of variables. The downward pointing arrow in **Figure 2** shows the window's movement. The red color shows the convolutional filter used to extract features from the time series data. This filter will be shifted along the data window by a specific interval.

The kernel or filter used in convolution always has the same width as the time series (following the feature data), and the length can vary. In the convolution process, the kernel moves in one direction from the beginning of the first time series to the end [14]. The advantage of using CNN to extract features on univariate time series datasets is that it can recognize local patterns or features hidden in the data and convert them into more understandable parts of the model. This results in computational efficiency and the model's accuracy. Although most CNN applications consider non-temporal image data, this study expands its realm to temporal data effectively for time series data forecasting by teaming up with LSTM.

Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks turn out to be a powerful duo for predicting $PM_{2.5}$ levels. On the one hand, CNN reveals spatial patterns, helping us understand where pollution is coming from and how it spreads locally. On the other hand, LSTM figures out how pollution levels change over time—whether it's daily, seasonally, or over the long term. LSTM is also effective in handling data collected in irregular time intervals. This is another advantageous feature of LSTM in the prediction because data monitoring may not always be conducted on a strict schedule. By combining both their strengths, we can a smart system that can grasp the full picture of what's going on with $PM_{2.5}$, giving us better predictions and a clearer understanding of air quality.

C. CNN-LSTM

CNN works to extract knowledge in the representation of time series data, while LSTM identifies short-term and long-term dependencies [13]. One of CNN's main advantages is the local perception feature and weight sharing, which can significantly reduce the number of parameters and thus improve efficiency in the training process.

This CNN process consists of two main components: the convolutional layer (1D Conv) and the pooling layer. Each convolutional layer contains several convolution kernels, with

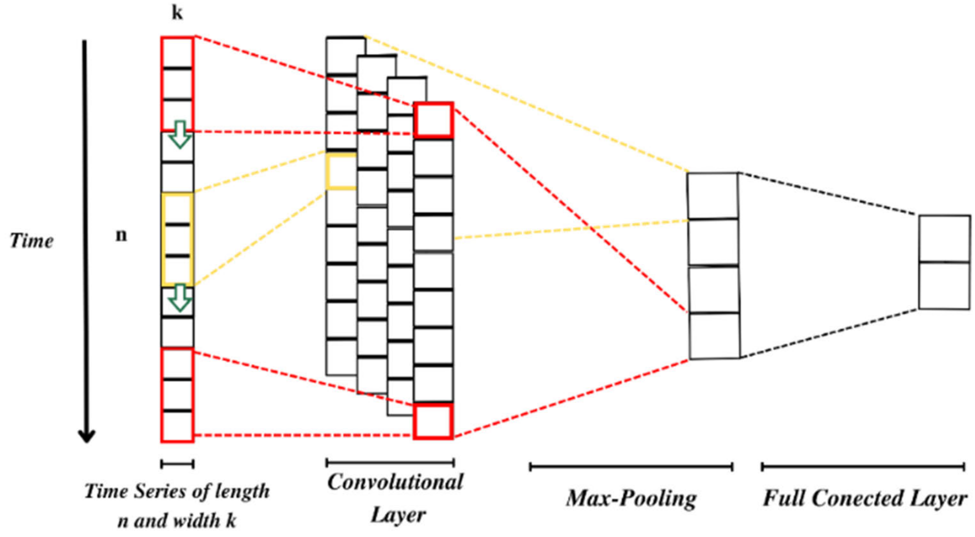


FIGURE 2. Our approches CNN for time series application.

calculations such as equation 5 below:

$$l_t = \tanh(x_t * W_t + b_t) \quad (5)$$

where l_t represents the output value of the convolution process, *tanh* is the activation function, x_t represents the input, W_t represents the weight of the convolution kernel, and b_t is the bias of the convolution kernel. After the convolution operation in the convolutional layer, the important features of the data are extracted, causing an increase in the feature dimension [90], [91]. Furthermore, there is a pooling layer to overcome the increase in feature dimension by reducing the number of extracted features again.

The following **Figure 3** illustrates the CNN-LSTM architecture model. The CNN process consists of 2 main components: the convolutional layer (1D Conv) that receives input from time steps in a 1D (one-dimensional) array. It then processes mathematical operations in extracting input data features by taking special features such as trends, patterns, or certain variations from $PM_{2.5}$ concentration data, with convolution operations followed by activation operations, such as ReLU, to add non-linearity to the output.

During this process, a feature representation matrix consists of several layers that represent feature extraction results from different filters (See yellow color). In the second layer of CNN, a pooling operation is performed using the Max-Pooling layer to reduce the input dimensions from the convolutional layer process resulting in smaller segments (See red color); dimensional reduction is made by selecting the maximum value of each piece to speed up the training process. Then, to process the data into the format needed by LSTM, there is a flattened layer changing the output of the CNN layer in the form of a matrix into a one-dimensional vector.

Followed by the LSTM process and Fully-connected Layer (FC), or dense layer, which helps take the output of LSTM

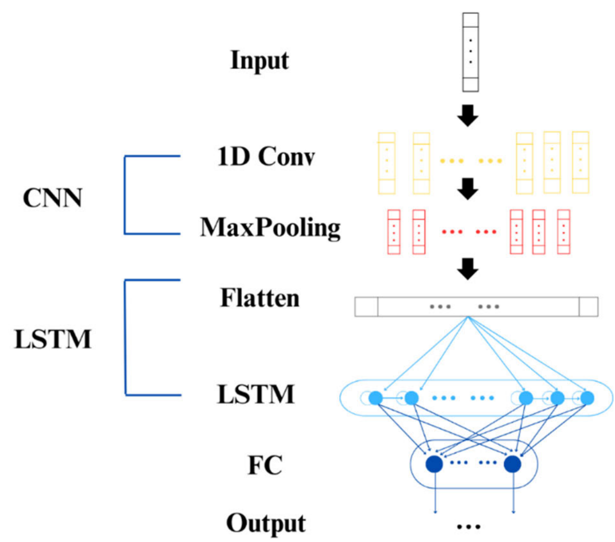


FIGURE 3. Our CNN-LSTM approch for time series application.

and process it into a predictive value [13]. Therefore, it can be explained that the work generated from the primary component or CNN layer will be collected to a smaller dimension and then channelled into the LSTM layer so that the output layer results in the form of predictions [17].

D. CONV-LSTM

CONV-LSTM is a one-dimensional convolutional model which contains convolution operations in LSTM cells [92]. This model can, then, process long-term dependencies. When the input matrix multiplication is calculated with LSTM cells, the process will be added with the convolution operation. The convolution operation takes two inputs, namely the kernel matrix and the input matrix. The kernel matrix scans the input matrix by multiplying each kernel element

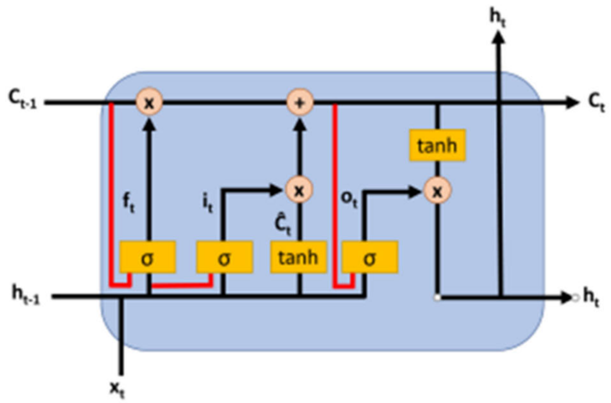


FIGURE 4. Conv-LSTM [93], [94], [95], [96].

with the corresponding component of the input and summing them [93], [94], [95]. The kernel weights are iteratively adjusted during training to optimize the prediction [96]. The CONV-LSTM cell has the same architecture as the LSTM, which consists of input gates, forget gates, output gates, and candidate values. In the CONV-LSTM cell, the input, forget, and output gate information is calculated using convolution operations on the hidden state and memory cells from the previous timestep h_{t-1} and the input at the current timestep X_t.

$$f_t = \sigma\left(W_{xf} * x_t + w_{hf} * h_{t-1} + W_{cf}^{\circ} c_{t-1} + b_f\right) \tag{6}$$

$$i_t = \sigma\left(W_{xi} * x_t + w_{hi} * h_{t-1} + W_{ci}^{\circ} c_{t-1} + b_i\right) \tag{7}$$

$$C_t = f_t^{\circ} C_{t-1} + i_t^{\circ} R(W_{xc} * x_t + W_{hc} * h_{t-1} + b_c \tag{8}$$

$$O_t = \sigma(W_{xo} * x_t + w_{ho} * h_{t-1} + W_{co}^{\circ} c_{t-1} + b_o) \tag{9}$$

$W_{cf}, W_{ci}, W_{co}, W_{hi}, W_{xi}, W_{ho}, W_{xo}, W_{xf}, W_{xc}, W_{hf}$ represent convolutional kernels used in the model, and b_i, b_f, b_o, b_C are bias vectors [92], [97]. **Figure 4** shows the CONV-LSTM architecture, where the red line indicates the additional connections found in the CONV-LSTM cell above the LSTM cell, which are derived from the current and previous cell states. The red line explains which forget gates, input gates, and output gates have a kernel matrix multiplication operation with the previous cell states $W_{cf}^{\circ} c_{t-1}$ (for instance, in forget gates). In addition to the LSTM's ability to capture temporal correlation and simultaneously represent detailed local information in the feature data by convolution process [82], [98], [99]. CONV-LSTM can help reduce the model size, especially for large input sizes. So the benefit of the CONV-LSTM method is that while the LSTM prior works well in terms of overall information interaction in weight calculation and convolution is more adaptable to represent more detailed local information.

III. DISCUSSION

The data used in this study is hourly observation data obtained from the Central Meteorology, Climatology and Geophysics Agency in 2022 on 21 April to 21 June regarding $PM_{2.5}$concentrations in the Kemayoran area, Central Jakarta. The data used in this study were 1488 data.

It was identified that there was an anomaly problem in the $PM_{2.5}$ data in the Kemayoran area, where data anomalies deviated from the observations of $PM_{2.5}$concentrations, which could be caused by errors in the equipment, such as an inadequate maintenance process. Due to anomalies, it can affect the results of the analysis. Also, the evaluation of the model to be produced, so in this study, the anomalies in the observation data will be removed as handling, which causes missing data.

In filling in the empty values, imputation of data is carried out, one of which is the interpolation process, which estimates unknown data points between two known issues. In this study, the interpolation method used is spline interpolation which has the advantage of being able to produce more minor errors and produce smoother interpolation results.

The datasets we use have values ranging from 1 to 91 $\mu g/m^3$ with an average value of 22.99 $\mu g/m^3$ yang which shows the concentration of $PM_{2.5}$ in Kemayoran is in the moderate category. Still, there are some observation data in certain time ranges that reach the unhealthy category (66-150 μg/m3) so it can be said that $PM_{2.5}$ concentrations can change at different times. Data scaling by equation (10) helps maintain the range of $PM_{2.5}$ concentration values so that they remain balanced for the performance improvement in training the datasets.

$$z = \frac{x - \mu}{\sigma} \tag{10}$$

The next step is to split data into three parts: training data, validation data, and test data. Data splitting involves determining the data by date to make it easier to read the comparison chart. In this study, 1488 $PM_{2.5}$ concentration data were split into several scenarios to improve the accuracy and generalization of the data described in **Figure 5.**

IV. RESULTS

This study employs the Tensorflow Keras library in conjunction with the Python programming language for analysis and predictive modeling. For additional details, a GitHub link is provided in the data acknowledgment section as a reference.

Table 1 shows the CNN-LSTM model. The first layer or layer is the input of the Convolutional Neural Network (CNN) architecture, where this 1D convolution layer functions in extracting features in the time series s. This layer is formed with three dimensions; the first dimension (None) is a sample (many rows of data) or the amount of input data used in a batch (batch size) and has not been determined during the model compilation process; the second dimension represents the time step used in prediction which is 24, and the last dimension represents the number of filters in the convolution process of 64.

Then, the next layer has an LSTM layer with 16 neuron units to process sequential data and produce output at each time st. The dense layer with 24 neuron units shows the

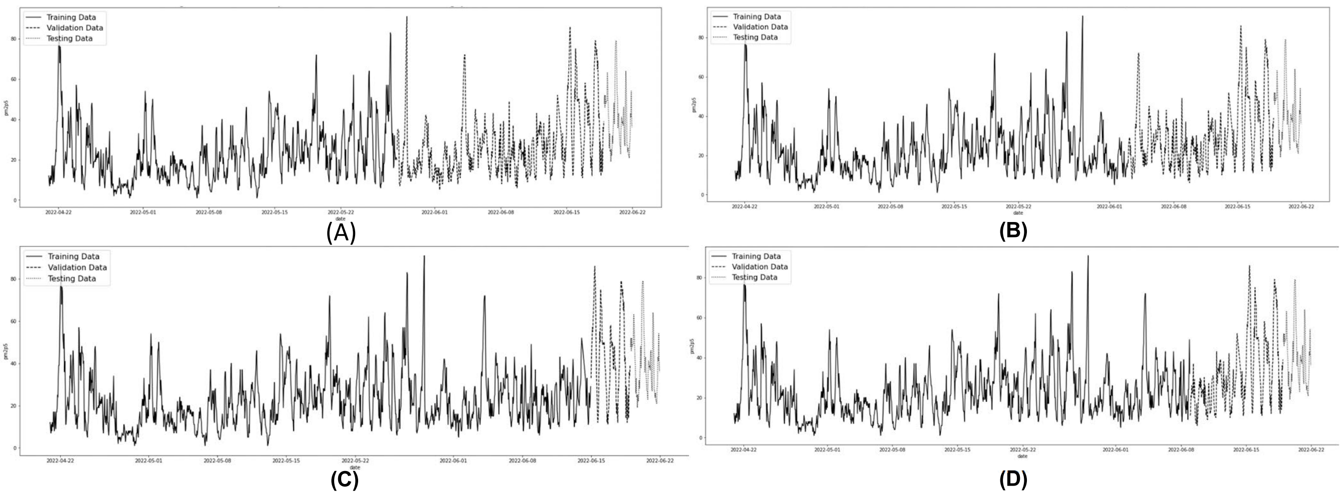


FIGURE 5. PM2.5 concentration by splitting 60:40 (A), 70:30 (B), 80:20(C), and 90:10 (D).

TABLE 1. Layer model of CNN-LSTM.

Layer (Type)	Output Shape	Parameter
Conv1d_1 (Conv1D)	(None,24,64)	256
Max_pooling1d(MaxPooling1D)	(None,24,64)	0
Lstm_2 (LSTM)	(None,24,16)	5184
Lstm_3(LSTM)	(None,16)	2112
Dense_1 (Dense)	(None,24)	408
Dense_2(Dense)	(None,24)	600

TABLE 2. Parameter setting of CONV-LSTM.

Layer	Output Shape	Parameter
CONV-LSTM_5	(None,24,32)	128
Bidirectional (Bidirectional)	(None,24,16)	2624
Bidirectional_1 (Bidirectional)	(None,16)	1600
Dense_2 (Dense)	(None,24)	408

coating has one output value, namely the prediction target value for the next 24 hours.

The CNN-LSTM model has demonstrated proficiency in extracting spatial features and local patterns from the $PM_{2.5}$ data, particularly good at capturing the distribution of pollutants. This capability allows for a nuanced understanding of localized pollution sources and the spatial dynamics influencing $PM_{2.5}$ concentrations. On the other hand, the CONV-LSTM model, with its integrated convolutional and LSTM layers, has exceled in simultaneously capturing both spatial and temporal dependencies in the time series data. The convolution operations enhance feature extraction, while LSTM handles long-term temporal dependencies, providing a comprehensive approach to modeling the intricate patterns within the $PM_{2.5}$ concentration data. Through the performance assessments of these models, this work underscores the significance of a hybrid approach that combines spatial and temporal modeling. The insights gained contribute to the optimization of $PM_{2.5}$ forecasting models, guiding future research in selecting or adapting hybrid architectures based on specific data characteristics and objectives.

Table 2 is a CONV-LSTM model with layers that almost resemble CNN-LSTM, where the difference between these two models is only in the combination of convolution and LSTM layers. In the CONV-LSTM model, the LSTM architecture used is bidirectional concerning previous research, wherein [100] predicted $PM_{2.5}$ in Beijing using a hybrid model, namely CONV-LSTM using bi-LSTM architecture, to focus on studying the temporal correlation in $PM_{2.5}$ concentrations. The parameters used in the compiling process include using an optimizer with the Adam algorithm, the learning rate initiated is 1e-5, and for the activation, the function used ReLU with the loss function chosen is Huber loss. The training process of CNN-LSTM and CONV-LSTM models used an optimal epoch of 250 iterations. In addition, using several testing schemes, including comparing the number of neurons and batch size set at 32, applying regularizers, and the number of LSTM layers used in both models. After obtaining the optimal CNN-LSTM and CONV-LSTM models, the following is a loss graph from the training and validation process of both models based on four data-splitting scenarios:

The results of CNN-LSTM and CONV-LSTM modeling testing obtained the results of the loss graph as in **Figure 6** shows that in the training process of the two models for each scenario, the loss value decreases in each iteration until a stable point. There is no indication of overfitting because there is no gap or considerable distance between the training and validation loss values.

We use three evaluation metrics: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE), for the assessment of predictive model performance, which are defined as Equation 10 to 12 where n as the number of observation, Y_i is the actual value and $\hat{Y}_i$ is the predicted value.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|Y_i - \hat{Y}_i\right| \tag{11}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2} \tag{12}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{\left|Y_i - \hat{Y}_i\right|}{Y} \times 100\% \tag{13}$$

MAE serves as a direct measure, offering a clear gauge of accuracy by assessing the average magnitude of errors between predicted and actual values. RMSE introduces a nuanced perspective by considering the square of errors, which adds more weights to large errors. MAPE is a relative measure, expressing the percentage difference between predicted and actual values. This provides a valuable insight into the proportional accuracy of the model. The synergistic use of these metrics not only ensures a precise evaluation of accuracy but also facilitates effective comparisons across various models. This holistic evaluation is pivotal in guiding the refinement and optimization of predictive models, empowering researchers and practitioners to make well-informed decisions regarding the suitability and effectiveness of their models.

Based on the test conducted, **Table 3** shows the results of the model performance evaluation. Predicting $PM_{2.5}$ concentrations using the CONV-LSTM method with scenario 4, namely the 90:10 ratio data splitting, dominates the better accuracy and efficiency of error values compared to the CNN-LSTM method. Suppose we refer to one of the test data metric evaluations, which is unseen data or data that has never been seen by the model, namely the MAPE Test parameter, in testing the accuracy and feasibility of the model. In that case, CONV-LSTM has a more efficient model in making predictions with a MAPE value of 16.39% when compared to CNN-LSTM, with a MAPE value of 17.92%. Therefore, CONV-LSTM is the model that will be used in predicting $PM_{2.5}$ concentrations in the Kemayoran area, Central Jakarta.

Categorizing MAPE values as "good enough" holds profound significance, signifying the satisfactory performance of the predictive model for practical applications. When MAPE values attain the "good enough" classification, it assures that, on average, the model's predictions align acceptably with actual values. The contextual importance of these values is further emphasized through their comparison to industry

TABLE 3. Metric evaluation.

Metode	Scenario		MAE	RMSE	MAPE
CNN-LSTM	**1**	**Training**	5.02	6.9	25.20%
	2		5.58	7.29	25%
	3		5.51	7.37	33.40%
	4		**4.94**	**6.83**	**23.91%**
	1	**Validation**	5.87	7.9	23.69%
	2		6.84	8.59	22.49%
	3		6.87	8.67	22.31%
	4		**6.41**	**8.12**	**17.86%**
	1	**Testing**	7.33	9.46	17.95%
	2		7.92	9.7	18.66%
	3		7.75	9.4	18.60%
	4		**7.35**	**9.32**	**17.92%**
Metode	**Scenario**		**MAE**	**RMSE**	**MAPE**
CONV-LSTM	**1**	**Training**	**6.67**	**7.93**	**28.40%**
	2		**4.78**	6.88	24.82%
	3		**4.73**	6.56	23.64%
	4		4.7	**6.5**	**23.17%**
	1	**Validation**	**7.32**	**9.09**	**25.37%**
	2		**5.95**	7.88	$2_{2.5}5$%
	3		**5.54**	7.55	20.41%
	4		5.6	**7.38**	**18.18%**
	1	**Testing**	7.35	9.32	17.92%
	2		**8.47**	9.96	19.25%
	3		**6.97**	9.12	17.83%
	4		**6.84**	8.83	17.14%

standards or guidelines. If the obtained MAPE values meet or surpass established benchmarks, it signals that the model aligns with industry expectations. However, in the absence of specific benchmarks from BMKG Indonesia to label forecasting results as "good," it underscores the need for a nuanced evaluation and consideration of industry-specific precision requirements. In essence, the categorization of MAPE values serves as a valuable indicator of the model's readiness for practical deployment and decision-making, acknowledging the current absence of predefined benchmarks from the relevant authority.

After obtaining the best model based on metric evaluation, forecasting is performed 24 hours for data on 22 June 2022 from 00:00 to 23:00. In the future using the CONV-LSTM model. The forecasting results generated from the CONV-LSTM model will then be descaled from the inverse z-score based on equation 8. Here are the results of $PM_{2.5}$

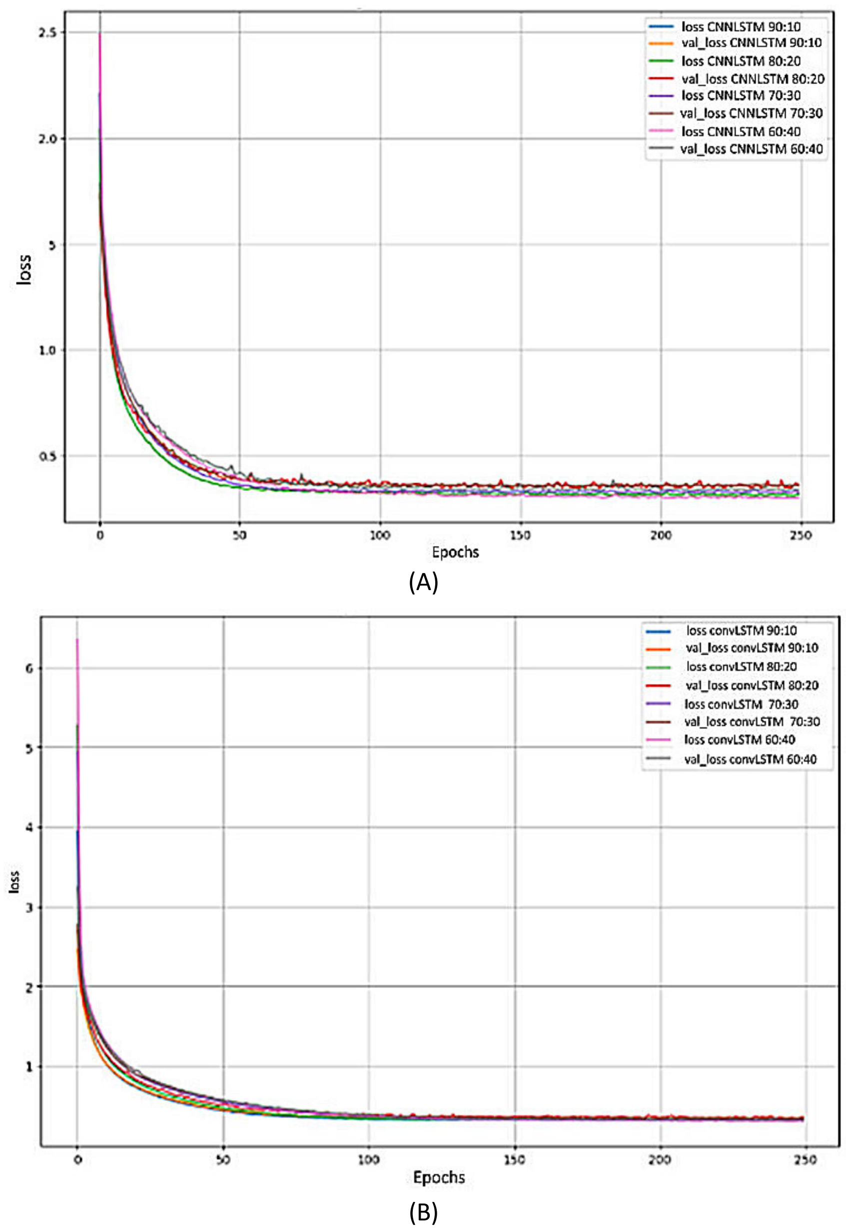


FIGURE 6. Training & validation loss CNNLSTM (A) and CONV-LSTM (B).

concentration forecasting on 22 June 2022 from 00:00 to 23:00.

The results of forecasting $PM_{2.5}$concentrations on 22 June 2022 have the highest levels of $PM_{2.5}$concentrations worth 34.87 μg/m^3 in the early morning, where in the early morning there are weather changes such as a decrease in temperature and high humidity at night. The average $PM_{2.5}$concentration on the 22nd was 27.37 μg/m^3, which shows that the air quality in Kemayoran is still quite good, at a moderate level. The prediction results also show differences during the day and night, so a Diurnal pattern is identified.

Implementing data science in SDGs policy can directly or indirectly focus data on becoming accurate information with technological methods as automated as needed. Achieving the 17 SDG goals requires support from all levels of society with various disciplines and knowledge, especially data scientists and collaborating academics/students. Besides, the executive, legislative, and judiciary support makes the position of science data can take its best place in contributing to sustainable development.

Data science has become a tool to synergize the 17 SDG goals as a form of implementation of sustainable and equitable development in Indonesia. The recommendations that need to be followed up are that achieving SDGs in the making various strategic decisions/policies is inseparable from the role of data and information processed by Data

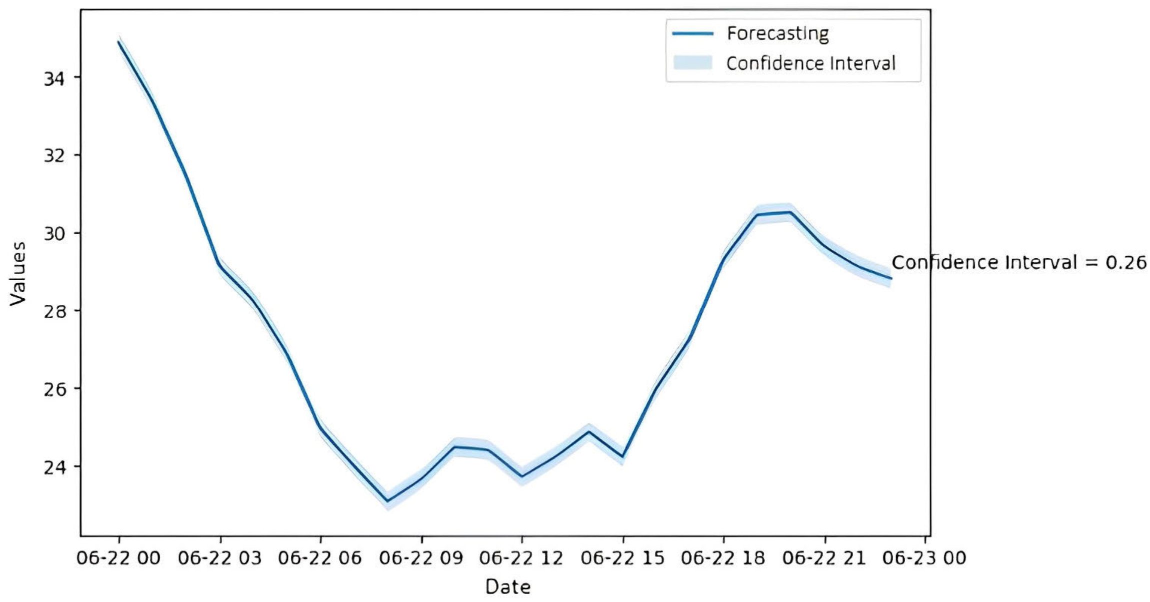


FIGURE 7. Nowcasting $PM_{2.5}$ concentration 24 hours ahead on 22 June 2022 with confidence interval.

Scientists, which requires coordination, cooperation, synergy, and partnership from each unit of related institutions in fulfilling the goals and apologies of science data in achieving SDGs in the Republic of Indonesia and establishing data science as the spearhead in decision making of stakeholders in the fulfillment and achievement of SDGs.

At its core, this research plays a pivotal role in the accurate prediction of pollution indicators, serving as a robust early warning system. This capability empowers authorities to anticipate air pollution and optimize waste management strategies proactively. The broader significance of our work is clearly seen by its seamless alignment with Sustainable Development Goal (SDG) indicator 11.6, aiming to curtail adverse per capita environmental impacts through advancements in air quality and waste management practices. In line with the mission of BMKG (Badan Meteorologi, Klimatologi, dan Geofisika), tasked with executing governmental responsibilities in meteorology, climatology, air quality, and geophysics, our research stands as a valuable tool. Its outcomes can substantially assist in fulfilling these duties, offering practical insights for the effective management of pollution.

The hybrid model, seamlessly merging Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks, serves as a critical advancement in bolstering early warning systems for air pollution. Specifically tailored for the vibrant landscape of Kemayoran, DKI Jakarta, this model excels in capturing the intricacies of the region's air quality dynamics. The CNN component's acute spatial sensitivity allows for precise identification of localized pollution sources and complex spatial patterns inherent in a busy urban environment. Simultaneously, the LSTM component adeptly navigates the temporal intricacies of $PM_{2.5}$ concentration time series data, crucial for anticipating pollution fluctuations in a dynamic area like Kemayoran. By integrating both spatial and temporal aspects, the hybrid model provides a comprehensive analysis, enhancing the accuracy of early warnings. Its adaptability to the ever-changing conditions of Kemayoran, coupled with the capacity for customization to local factors, positions this model as an invaluable tool. Ultimately, the timely and precise predictions furnished by the hybrid model empower authorities to make informed decisions, implement preventive measures, and optimize strategies tailored to Kemayoran's unique air quality challenges.

V. PRACTICAL IMPLICATION

Based on the findings described in the previous section, we provide practical recommendations for predictive modeling as a recommendation. Initially, ensure the data is well-organized and consistent by identifying anomalies in the observation data. Following this, employ spline interpolation to fill in any missing data, thereby enhancing the overall structure and completeness of the dataset.

Also scale the data with z-score so that it can help keep the range of values of each $PM_{2.5}$ concentration in the dataset balanced where it is not too large or small. Perform data splitting with various ratio scenarios like 60:40, 70:30, 80:20, 90:10 to find the ideal number of splits for each algorithm in machine learning. It is essential to perform hyperparameter building and testing both models to produce good predictions. The model we employed for forecasting $PM_{2.5}$ concentration is CONV-LSTM, and we're comparing it to the CNN-LSRTM predictive method using the smallest metric evaluation measurement set available.

The CONV-LSTM model emerges as a robust solution for real-time $PM_{2.5}$ nowcasting, offering a seamless integration of spatial and temporal information. This model's unique combination of convolutional and LSTM layers showcases exceptional proficiency in feature extraction, unraveling intricate spatial patterns within the $PM_{2.5}$ dataset. Notably, the CONV-LSTM model excels in handling long-term temporal dependencies, providing a nuanced understanding of historical trends critical for precise nowcasting. Its predictive accuracy surpasses that of the CNN-LSTM model, especially in scenarios where both spatial and temporal factors play significant roles in air quality dynamics. The model's adaptability to irregular time intervals in practical monitoring scenarios further solidifies its reliability. Ultimately, the CONV-LSTM model's holistic approach and comprehensive grasp of spatial-temporal dynamics put it as a formidable tool for advancing $PM_{2.5}$ nowcasting, particularly in regions characterized by diverse and dynamic pollution sources.

The study establishes a groundwork for future research in air quality modeling, inviting researchers to assess and compare various hybrid models. This approach enables the identification of the most effective strategies tailored to diverse urban contexts, thereby contributing to the continuous enhancement of air quality forecasting systems. The findings resonate with broader implications, offering potential advancements in accuracy, comprehension, and early warning capabilities within urban air quality monitoring and forecasting. The adoption of hybrid models, exemplified in the study, emerges as a promising direction for navigating the intricacies of urban air quality dynamics, promising substantial improvements in pollution management effectiveness.

VI. CONCLUSION

This work has shown that both models, CNN-LSTM and CONV-LSTM, are suitable for predicting $PM_{2.5}$ concentrations. At the time of splitting the data with several different ratios, the division by 90:10 is the best ratio for both CNN-LSTM and CONV-LSTM predictive models, with a large enough training dataset for training various possible patterns in $PM_{2.5}$ concentration data and also with validation set of sufficient size to evaluate the performance of the model quite well. Indeed, the majority of training sets were large enough to introduce the model to various possible patterns in $PM_{2.5}$ concentration data, and validation sets were also large enough to evaluate the model's performance.

The results by the CNN-LSTM model showed MAE worth 7.35, RMSE 9.32, and MAPE 17.92%. As for the CONV-LSTM model, we obtained MAE worth 6.52, RMSE 8.55 and MAPE 16.39%. While both models produced MAPE values that fall into the good enough range with values $<$20%, the CONV-LSTM model obtained overall better metric evaluation values.

However, this study has limitations in that it is known that many factors cause $PM_{2.5}$ concentrations, one of which is meteorological factors that are not considered in this study. For example, despite the fact that specific Kemayoran Jakarta regions are more prone to air pollution, this study does not consider meteorological impact on air quality in Kemayoran, Jakarta. Therefore, in future research, it is expected to forecast $PM_{2.5}$ concentrations using several possible relevant features, with the advantage of using convolutional neural networks that can extract spatial features, making it possible to use meteorological factors such as weather, wind speed, wind direction, etc. at several different regional observation points by modifying model hyperparameters such as the number of layers, number of neurons, learning rate and other parameters or can perform automatic hyperparameters in $PM_{2.5}$ prediction for real-time output with more efficient and optimal model performance results.

The methodology in this study demonstrates the efficacy of integrating data from diverse sources, including air quality observations and meteorological data. Subsequent advancements in research could investigate comparable data fusion techniques to forecast additional atmospheric pollutants, taking into account the accessibility and pertinence of a variety of datasets.

Furthermore, for future research, it is recommended to provide a detailed exploration of the architectural distinctions between the CNN-LSTM and CONV-LSTM models, specifically within the context of predicting $PM_{2.5}$ concentrations. Delving into the nuances of these models' architectures will contribute to a deeper understanding of their individual strengths and weaknesses, enabling a more comprehensive evaluation of their performance in air quality prediction. This exploration could shed light on the specific features or patterns each architecture excels at capturing, offering valuable insights for refining and optimizing predictive models in subsequent studies.

The superior performance of the CONV-LSTM model in $PM_{2.5}$ nowcasting opens up avenues for practical applications in operational forecasting, early warning systems, and decision support. Customization for specific urban environments, exploration of transferability to other pollutants, and collaboration for further validation contribute to the model's real-world impact on air quality management.

COMPETING INTERESTS

The authors declare no competing interests.

DATA AVAILABILITY

The source code and the material and findings data of this study are openly available in full access by the corresponding author.

DECLARATION OF COMPETING INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

AUTHOR CONTRIBUTION

Tafia Hasna Putri and Rezzy Eko Caraka conceived the research and constructed the experimental design. Rezzy Eko

Caraka, Toni Toharudi, Yunho Kim, Rung-Ching Chen, Prana Ugiana Gio, Anjar Dimara Sakti, Resa Septiani Pontoh, and Bens Pardamean managed the project. Tafia Hasna Putri and Rezzy Eko Caraka analyzed the data. Rezzy Eko Caraka participated in the verification and interpretation of data. Tafia Hasna Putri and Rezzy Eko Caraka drew the study design, carried out data management, and constructed a database. Tafia Hasna Putri and Rezzy Eko Caraka finalized the instrument. Tafia Hasna Putri, Rezzy Eko Caraka, and Yunho Kim wrote the final manuscript. Tafia Hasna Putri, Rezzy Eko Caraka, Toni Toharudi, Yunho Kim, Rung-Ching Chen, Prana Ugiana Gio, Anjar Dimara Sakti, Resa Septiani Pontoh, Indah Reski Pratiwi, Farid Azhar Lutfi Nugraha, Thalita Safa Azzahra, Jessica Jesslyn Cerelia, Gumgum Darmawan, Defi Yusti Faidah, and Bens Pardamean read and approved the final manuscript.

ACKNOWLEDGMENT

(Tafia Hasna Putri and Rezzy Eko Caraka are co-first authors.)

REFERENCES

[1] WHO. (2021). *Review of Evidence on Health Aspects of Air Pollution-REVIHAAP*. [Online]. Available: https://www.who.int/europe/publications/i/item/WHO-EURO-2013-4101-43860-61757

[2] J. Zhang, Y. Wei, and Z. Fang, "Ozone pollution: A major health hazard worldwide," *Frontiers Immunology*, vol. 10, pp. 1–10, Oct. 2019, doi: 10.3389/fimmu.2019.02518.

[3] L. Liao, M. Du, and Z. Chen, "Air pollution, health care use and medical costs: Evidence from China," *Energy Econ.*, vol. 95, Mar. 2021, Art. no. 105132, doi: 10.1016/j.eneco.2021.105132.

[4] R.-Y. Chen, K.-F. Ho, G.-B. Hong, and K.-J. Chuang, "Houseplant, indoor air pollution, and cardiovascular effects among elderly subjects in taipei, Taiwan," *Sci. Total Environ.*, vol. 705, Feb. 2020, Art. no. 135770, doi: 10.1016/j.scitotenv.2019.135770.

[5] H. Akimoto, "Global air quality and pollution," *Science*, vol. 302, no. 5651, pp. 1716–1719, Dec. 2003, doi: 10.1126/science.1092666.

[6] B. Brunekreef and S. T. Holgate, "Air pollution and health," *Lancet*, vol. 360, no. 9341, pp. 1233–1242, Oct. 2002, doi: 10.1016/S0140-6736(02)11274-8.

[7] N. Singh and S. C. Davar, "Noise pollution-sources, effects and control," *J. Hum. Ecol.*, vol. 16, no. 3, pp. 181–187, Nov. 2004, doi: 10.1080/09709274.2004.11905735.

[8] R. Tosepu, J. Gunawan, D. S. Effendy, L. O. A. I. Ahmad, H. Lestari, H. Bahar, and P. Asfian, "Correlation between weather and COVID-19 pandemic in Jakarta, Indonesia," *Sci. Total Environ.*, vol. 725, Jul. 2020, Art. no. 138436, doi: 10.1016/j.scitotenv.2020.138436.

[9] T. Firman, I. M. Surbakti, I. C. Idroes, and H. A. Simarmata, "Potential climate-change related vulnerabilities in Jakarta: Challenges and current status," *Habitat Int.*, vol. 35, no. 2, pp. 372–378, Apr. 2011.

[10] S. D. A. Kusumaningtyas, E. Aldrian, T. Wati, D. Atmoko, and S. Sunaryo, "The recent state of ambient air quality in Jakarta," *Aerosol Air Quality Res.*, vol. 18, no. 9, pp. 2343–2354, 2018, doi: 10.4209/aaqr.2017.10.0391.

[11] T. Toharudin, R. E. Caraka, I. R. Pratiwi, Y. Kim, P. U. Gio, A. D. Sakti, M. Noh F. A. L. Nugraha, R. S. Pontoh, T. H. Putri, T. S. Azzahra, J. J. Cerelia, G. Darmawan, and B. Pardamean, "Boosting algorithm to handle unbalanced classification of $PM_{2.5}$ concentration levels by observing meteorological parameters in Jakarta-Indonesia using AdaBoost, XGBoost, CatBoost, and LightGBM," *IEEE Access*, vol. 11, pp. 35680–35696, 2023, doi: 10.1109/ACCESS.2023.3265019.

[12] R. E. Caraka, M. Noh, R.-C. Chen, Y. Lee, P. U. Gio, and B. Pardamean, "Connecting climate and communicable disease to penta helix using hierarchical likelihood structural equation modelling," *Symmetry*, vol. 13, no. 4, p. 657, Apr. 2021.

[13] I. Calzada, "Democratising smart cities? penta-helix multistakeholder social innovation framework," *Smart Cities*, vol. 3, no. 4, pp. 1145–1172, Oct. 2020, doi: 10.3390/smartcities3040057.

[14] J. Rycroft-Malone, J. E. Wilkinson, C. R. Burton, G. Andrews, S. Ariss, R. Baker, S. Dopson, I. Graham, G. Harvey, G. Martin, B. G. McCormack, S. Staniszewska, and C. Thompson, "Implementing health research through academic and clinical partnerships: A realistic evaluation of the collaborations for leadership in applied health research and care (CLAHRC)," *Implement. Sci.*, vol. 6, no. 1, pp. 1–2, Dec. 2011, doi: 10.1186/1748-5908-6-74.

[15] X. Gao and W. Li, "A graph-based LSTM model for $PM_{2.5}$ forecasting," *Atmos. Pollut. Res.*, vol. 12, no. 9, Sep. 2021, Art. no. 101150, doi: 10.1016/j.apr.2021.101150.

[16] N. Sobanapuram Muruganandam and U. Arumugam, "Dynamic ensemble multivariate time series forecasting model for $PM_{2.5}$," *Comput. Syst. Sci. Eng.*, vol. 44, no. 2, pp. 979–989, 2023, doi: 10.32604/csse.2023.024943.

[17] M.-C. Yang and M. C. Chen, "Composite neural network: Theory and application to PM2.5 prediction," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 2, pp. 1311–1323, Feb. 2023, doi: 10.1109/TKDE.2021.3099135.

[18] R. E. Caraka, R. C. Chen, T. Toharudin, B. Pardamean, H. Yasin, and S. H. Wu, "Prediction of status particulate matter 2.5 using state Markov chain stochastic process and HYBRID VAR-NN-PSO," *IEEE Access*, vol. 7, pp. 161654–161665, 2019, doi: 10.1109/ACCESS.2019.2950439.

[19] J. G. De Gooijer and R. J. Hyndman, "25 years of time series forecasting," *Int. J. Forecasting*, vol. 22, no. 3, pp. 443–473, Jan. 2006, doi: 10.1016/j.ijforecast.2006.01.001.

[20] G.-Y. Lin, H.-W. Chen, B.-J. Chen, and Y.-C. Yang, "Characterization of temporal $PM_{2.5}$, nitrate, and sulfate using deep learning techniques," *Atmos. Pollut. Res.*, vol. 13, no. 1, Jan. 2022, Art. no. 101260, doi: 10.1016/j.apr.2021.101260.

[21] X. Yan, Z. Zang, N. Luo, D. Li, and Y. Guo, "Retrieval of real-time $PM_{2.5}$, temperature and humidity profiles from satellite and ground-based remote sensing data using advanced deep learning models," in *Proc. 41st Asian Conf. Remote Sens.*, 2020, pp. 1–9.

[22] M.-C. Yang and M. C. Chen, "PM2.5 forecasting using pre-trained components," in *Proc. IEEE Int. Conf. Big Data*, Dec. 2018, pp. 4488–4491, doi: 10.1109/BIGDATA.2018.8622559.

[23] Z. Zhang, X. Ma, and K. Yan, "A deep learning model for $PM_{2.5}$ concentration prediction," in *Proc. IEEE Intl. Conf. Dependable, Autonomic Secure Comput., Intl. Conf. Pervasive Intell. Comput., Intl. Conf. Cloud Big Data Comput., Intl. Conf. Cyber Sci. Technol. Congr. (DASC/PiCom/CBDCom/CyberSciTech)*, Oct. 2021, pp. 428–433, doi: 10.1109/DASC-PICom-CBDCom-CyberSciTech52372.2021.00078.

[24] J. Li, N. Wang, Z.-H. Wang, H. Li, C.-C. Chang, and H. Wang, "New secret sharing scheme based on faster R-CNNs image retrieval," *IEEE Access*, vol. 6, pp. 49348–49357, 2018, doi: 10.1109/ACCESS.2018.2821690.

[25] P. Sharma, Y. P. S. Berwal, and W. Ghai, "Performance analysis of deep learning CNN models for disease detection in plants using image segmentation," *Inf. Process. Agricult.*, vol. 7, no. 4, pp. 566–574, Dec. 2020, doi: 10.1016/j.inpa.2019.11.001.

[26] Z. Liu, J. Hu, L. Weng, and Y. Yang, "Rotated region based CNN for ship detection," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2017, pp. 900–904.

[27] P. Napoletano, F. Piccoli, and R. Schettini, "Anomaly detection in nanofibrous materials by CNN-based self-similarity," *Sensors*, vol. 18, no. 2, p. 209, Jan. 2018, doi: 10.3390/s18010209.

[28] M. Canizo, I. Triguero, A. Conde, and E. Onieva, "Multi-head CNN–RNN for multi-time series anomaly detection: An industrial case study," *Neurocomputing*, vol. 363, pp. 246–260, Oct. 2019, doi: 10.1016/j.neucom.2019.07.034.

[29] W. Alakwaa, M. Nassef, and A. Badr, "Lung cancer detection and classification with 3D convolutional neural network (3D-CNN)," *Int. J. Adv. Comput. Sci. Appl.*, vol. 8, no. 8, p. 409, 2017.

[30] K. Murata, M. Mito, D. Eguchi, Y. Mori, and M. Toyonaga, "A single filter CNN performance for basic shape classification," in *Proc. 9th Int. Conf. Awareness Sci. Technol.*, Sep. 2018, pp. 139–143.

[31] R. Ghosh, K. Ghosh, and S. Maitra, "Automatic detection and classification of diabetic retinopathy stages using CNN," in *Proc. 4th Int. Conf. Signal Process. Integr. Netw. (SPIN)*, Feb. 2017, pp. 550–554.

[32] M. S. Islam, F. A. Foysal, N. Neehal, E. Karim, and S. A. Hossain, "InceptB: A CNN based classification approach for recognizing traditional Bengali games," *Proc. Comput. Sci.*, vol. 143, pp. 595–602, Jan. 2018.

[33] A. A. Khodabakhsh, "Forecasting multivariate time-series data using LSTM and mini-batches," in *Proc. 7th Int. Conf. Contemp. Issues Data Sci.*, Cham: Springer, 2019, pp. 121–129.

[34] N. Xue, I. Triguero, G. P. Figueredo, and D. Landa-Silva, "Evolving deep CNN-LSTMs for inventory time series prediction," in *Proc. IEEE Congr. Evol. Comput. (CEC)*, Jun. 2019, pp. 1517–1524, doi: 10.1109/CEC.2019.8789957.

[35] S. Mehtab, J. Sen, and S. Dasgupta, "Robust analysis of stock price time series using CNN and LSTM-based deep learning models," in *Proc. 4th Int. Conf. Electron., Commun. Aerosp. Technol. (ICECA)*, Nov. 2020, pp. 1481–1486, doi: 10.1109/ICECA49313.2020.9297652.

[36] S. Mehtab and J. Sen, "Analysis and forecasting of financial time series using CNN and LSTM-based deep learning models," in *Proc. Adv. Distrib. Comput. Mach. Learn. Proc. (ICADCML)*, vol. 302, 2022, pp. 405–423, doi: 10.1007/978-981-16-4807-6_39.

[37] L. Wang, Y. Zhang, K. Wang, B. Zheng, Q. Zhang, and W. Wei, "Application of weather research and forecasting model with chemistry (WRF/Chem) over northern China: Sensitivity study, comparative evaluation, and policy implications," *Atmos. Environ.*, vol. 124, pp. 337–350, Jan. 2016, doi: 10.1016/j.atmosenv.2014.12.052.

[38] N. Masseran and M. A. M. Safari, "Modeling the transition behaviors of PM_{10} pollution index," *Environ. Monitor. Assessment*, vol. 192, no. 7, pp. 1–15, Jul. 2020, doi: 10.1007/s10661-020-08376-1.

[39] N. Masseran and M. A. M. Safari, "Risk assessment of extreme air pollution based on partial duration series: IDF approach," *Stochastic Environ. Res. Risk Assessment*, vol. 34, nos. 3–4, pp. 545–559, Apr. 2020, doi: 10.1007/s00477-020-01784-2.

[40] N. Masseran and M. A. Mohd Safari, "Intensity–duration–frequency approach for risk assessment of air pollution events," *J. Environ. Manage.*, vol. 264, Jun. 2020, Art. no. 110429, doi: 10.1016/j.jenvman.2020.110429.

[41] K. Zhang, X. Yang, H. Cao, J. Thé, Z. Tan, and H. Yu, "Multi-step forecast of $PM_{2.5}$ and PM_{10} concentrations using convolutional neural network integrated with spatial–temporal attention and residual learning," *Environ. Int.*, vol. 171, Jan. 2023, Art. no. 107691, doi: 10.1016/j.envint.2022.107691.

[42] R. G. Brereton and G. R. Lloyd, "Support vector machines for classification and regression," *Analyst*, vol. 135, no. 2, pp. 230–267, 2010, doi: 10.1039/b918972f.

[43] V. J. Schwanitz, F. Piontek, C. Bertram, and G. Luderer, "Long-term climate policy implications of phasing out fossil fuel subsidies," *Energy Policy*, vol. 67, pp. 882–894, Apr. 2014, doi: 10.1016/j.enpol.2013.12.015.

[44] J. B. Gao, S. R. Gunn, C. J. Harris, and M. Brown, "A probabilistic framework for SVM regression and error bar estimation," *Mach. Learn.*, vol. 46, pp. 71–89, 2002.

[45] L. Breiman, "Random forest," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001, doi: 10.1023/a:1010933404324.

[46] C. Strobl, A. L. Boulesteix, T. Kneib, T. Augustin, and A. Zeileis, "Conditional variable importance for random forests," *BMC Bioinf.*, vol. 1, pp. 1–11, Dec. 2008, doi: 10.1186/1471-2105-9-307.

[47] D. R. Cutler, T. C. Edwards, K. H. Beard, A. Cutler, K. T. Hess, J. Gibson, and J. J. Lawler, "Random forests for classification in ecology," *Ecology*, vol. 88, no. 11, pp. 2783–2792, Nov. 2007, doi: 10.1890/07-0539.1.

[48] L. Breiman, *Consistency for a Simple Model of Random Forests*, document Tech. Rep. 670, Statist. Dept., Univ. California at Berkeley, 2004. [Online]. Available: https://www.stat.berkeley.edu/~breiman/RandomForests/consistencyRFA.pdf

[49] M. Ma, G. Zhao, B. He, Q. Li, H. Dong, S. Wang, and Z. Wang, "XGBoost-based method for flash flood risk assessment," *J. Hydrol.*, vol. 598, Jul. 2021, Art. no. 126382, doi: 10.1016/j.jhydrol.2021.126382.

[50] H. Dai, G. Huang, H. Zeng, and F. Zhou, "$PM_{2.5}$ volatility prediction by XGBoost-MLP based on GARCH models," *J. Cleaner Prod.*, vol. 356, Jul. 2022, Art. no. 131898, doi: 10.1016/j.jclepro.2022.131898.

[51] R. Mitchell, A. Adinets, T. Rao, and E. Frank, "XGBoost: Scalable GPU accelerated learning," 2018, *arXiv:1806.11248*.

[52] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2016, doi: 10.1145/2939672.2939785.

[53] H. Drucker, C. J. C. Burges, L. Kaufman, A. Smola, and V. Vapnik, "Support vector regression machines," in *Proc. Int. Conf. Adv. Neural Inf. Process. Syst.*, vol. 9, 1997, pp. 155–161.

[54] W. Chu and S. S. Keerthi, "Support vector ordinal regression," *Neural Comput.*, vol. 19, no. 3, pp. 792–815, Mar. 2007, doi: 10.1162/neco.2007.19.3.792.

[55] A. J. Smola and B. Schölkopf, "A tutorial on support vector regression," *Statist. Comput.*, vol. 14, no. 3, pp. 199–222, Aug. 2004.

[56] S. Jeya and L. Sankari, "Air pollution prediction by deep learning model," in *Proc. 4th Int. Conf. Intell. Comput. Control Syst. (ICICCS)*, May 2020, pp. 736–741, doi: 10.1109/ICICCS48265.2020.9120932.

[57] T. Xayasouk and H. Lee, "Air pollution prediction system using deep learning," *WIT Trans. Ecol. Environ.*, vol. 230, pp. 71–79, 2018, doi: 10.2495/AIR180071.

[58] J. L. Torres, A. García, M. De Blas, and A. De Francisco, "Forecast of hourly average wind speed with ARMA models in Navarre (Spain)," *Sol. Energy*, vol. 79, no. 1, pp. 65–77, Jul. 2005, doi: 10.1016/j.solener.2004.09.013.

[59] Y. Nie, "Computer intelligent value evaluation model through ARMA and long short-term memory neural network," in *Proc. 3rd Int. Conf. Artif. Intell. Adv. Manuf.*, Oct. 2021, pp. 2164–2169, doi: 10.1145/3495018.3501071.

[60] Y. Bai, Y. Li, B. Zeng, C. Li, and J. Zhang, "Hourly $PM_{2.5}$ concentration forecast using stacked autoencoder model with emphasis on seasonality," *J. Cleaner Prod.*, vol. 224, pp. 739–750, Jul. 2019, doi: 10.1016/j.jclepro.2019.03.253.

[61] K. K. R. Samal, K. S. Babu, S. K. Das, and A. Acharaya, "Time series based air pollution forecasting using SARIMA and prophet model," in *Proc. Int. Conf. Inf. Technol. Comput. Commun.*, Aug. 2019, pp. 80–85.

[62] M. Takruri, A. Abubakar, A.-H. Jallad, B. Altawil, P. R. Marpu, and A. Bermak, "Machine learning-based estimation of $PM_{2.5}$ concentration using ground surface DoFP polarimeters," *IEEE Access*, vol. 10, pp. 23489–23496, 2022, doi: 10.1109/ACCESS.2022.3151632.

[63] C. Minutti-Martinez, M. Arellano-Vàzquez, and M. Zamora-Machado, "A hybrid model for the prediction of air pollutants concentration, based on statistical and machine learning techniques," in *Proc. Mex. Int. Conf. Artif. Intell.*, vol. 13068, 2021, pp. 252–264, doi: 10.1007/978-3-030-89820-5_21.

[64] J. Song, K. Han, and M. E. J. Stettler, "Deep-MAPS: Machine-learning-based mobile air pollution sensing," *IEEE Internet Things J.*, vol. 8, no. 9, pp. 7649–7660, May 2021, doi: 10.1109/JIOT.2020.3041047.

[65] Y. Gu, B. Li, and Q. Meng, "Hybrid interpretable predictive machine learning model for air pollution prediction," *Neurocomputing*, vol. 468, pp. 123–136, Jan. 2022, doi: 10.1016/j.neucom.2021.09.051.

[66] S. Makridakis, E. Spiliotis, and V. Assimakopoulos, "M5 accuracy competition: Results, findings, and conclusions," *Int. J. Forecasting*, vol. 38, no. 4, pp. 1346–1364, Oct. 2022, doi: 10.1016/j.ijforecast.2021.11.013.

[67] S. Makridakis and M. Hibon, "The M3-competition: Results, conclusions and implications," *Int. J. Forecasting*, vol. 16, no. 4, pp. 451–476, Oct. 2000, doi: 10.1016/s0169-2070(00)00057-1.

[68] S. Makridakis, E. Spiliotis, and V. Assimakopoulos, "Predicting/hypothesizing the findings of the m4 competition," *Int. J. Forecasting*, vol. 36, no. 1, pp. 29–36, Jan. 2020, doi: 10.1016/j.ijforecast.2019.02.012.

[69] S. Makridakis, E. Spiliotis, and V. Assimakopoulos, "The M4 competition: 100,000 time series and 61 forecasting methods," *Int. J. Forecasting*, vol. 36, no. 1, pp. 54–74, Jan. 2020, doi: 10.1016/j.ijforecast.2019.04.014.

[70] A. Masood and K. Ahmad, "A model for particulate matter ($PM_{2.5}$) prediction for Delhi based on machine learning approaches," *Proc. Comput. Sci.*, vol. 167, pp. 2101–2110, 2020, doi: 10.1016/j.procs.2020.03.258.

[71] X. Li, L. Peng, X. Yao, S. Cui, Y. Hu, C. You, and T. Chi, "Long short-term memory neural network for air pollutant concentration predictions: Method development and evaluation," *Environ. Pollut.*, vol. 231, pp. 997–1004, Dec. 2017, doi: 10.1016/j.envpol.2017.08.114.

[72] W. Lu, J. Li, Y. Li, A. Sun, and J. Wang, "A CNN-LSTM-based model to forecast stock prices," *Complexity*, vol. 2020, pp. 1–10, Nov. 2020, doi: 10.1155/2020/6622927.

[73] J. Wang, X. Li, L. Jin, J. Li, Q. Sun, and H. Wang, "An air quality index prediction model based on CNN-ILSTM," *Sci. Rep.*, vol. 12, no. 1, p. 8373, May 2022, doi: 10.1038/s41598-022-12355-6.

[74] R. E. Caraka, R. C. Chen, H. Yasin, S. Suhartono, Y. Lee, and B. Pardamean, "Hybrid vector autoregression feedforward neural network with genetic algorithm model for forecasting space-time pollution data," *Indonesian J. Sci. Technol.*, vol. 6, no. 1, pp. 243–266, Jan. 2021.

[75] C. Chatfield, S. C. Wheelwright, and S. Makridakis, "Forecasting methods for management," *J. Roy. Stat. Society. Ser. A (General)*, vol. 141, no. 1, p. 113, 1978, doi: 10.2307/2344788.

[76] S. Makridakis, "A survey of time series," *Int. Stat. Rev. Revue Internationale de Statistique*, vol. 44, no. 1, p. 29, Apr. 1976, doi: 10.2307/1402964.

[77] E. Theodorou, S. Wang, Y. Kang, E. Spiliotis, S. Makridakis, and V. Assimakopoulos, "Exploring the representativeness of the M5 competition data," *Int. J. Forecasting*, vol. 38, no. 4, pp. 1500–1506, Oct. 2022, doi: 10.1016/j.ijforecast.2021.07.006.

[78] L. Wiranda and M. Sadikin, "Penerapan long short term memory pada data time series untuk memprediksi penjualan produk pt. metiska farma," Ph.D. Dissertation, Dept. Inform., Universitas Mercu Buana Jakarta, West Jakarta, Indonesia, 2019.

[79] T. Toharudin, R. S. Pontoh, R. E. Caraka, S. Zahroh, Y. Lee, and R. C. Chen, "Employing long short-term memory and Facebook prophet model in air temperature forecasting," *Commun. Statist. Simul. Comput.*, vol. 52, no. 2, pp. 279–290, Feb. 2023, doi: 10.1080/03610918.2020.1854302.

[80] Z. He, J. Zhou, H.-N. Dai, and H. Wang, "Gold price forecast based on LSTM-CNN model," in *Proc. IEEE Intl. Conf. Dependable, Autonomic Secure Comput., Intl. Conf. Pervasive Intell. Comput., Intl. Conf. Cloud Big Data Comput., Intl. Conf. Cyber Sci. Technol. Congr. (DASC/PiCom/CBDCom/CyberSciTech)*, Aug. 2019, pp. 1046–1053, doi: 10.1109/DASC/PiCom/CBDCom/CYBERSCITECH.2019.00188.

[81] K. Johan, J. C. Young, and S. Hansun, "LSTM-RNN automotive stock price prediction," *Int. Journal Sci. Technol. Res.*, vol. 8, no. 9, pp. 173–176, 2019.

[82] W. Yin and H. Schütze, "Attentive convolution: Equipping CNNs with RNN-style attention mechanisms," *Trans. Assoc. Comput. Linguistics*, vol. 6, pp. 687–702, Dec. 2018, doi: 10.1162/tacl_a_00249.

[83] A. H. Rahimyar, H. Q. Nguyen, and X. Wang, "Stock forecasting using M-Band wavelet-based SVR and RNN-LSTMs models," in *Proc. 2nd Int. Conf. Inf. Syst. Comput. Aided Educ. (ICISCAE)*, Sep. 2019, pp. 234–240, doi: 10.1109/ICISCAE48440.2019.221625.

[84] S. Sequeira and P. K. N. Banu, "Comparisons of stock price predictions using stacked RNN-LSTM," in *Proc. Data Sci. Comput. Intell., 16th Int. Conf. Inf. Process.*, vol. 1483, 2021, pp. 380–390, doi: 10.1007/978-3-030-91244-4_30.

[85] Colah. (2015). *Understanding LSTM Networks*. [Online]. Available: http://colah.github.io/posts/2015-08-Understanding-LSTMs/

[86] S. Geetha and L. Prasika, "Ground level ozone prediction for Delhi using LSTM-RNN," *Int. J. Innov. Technol. Exploring Eng.*, vol. 8, no. 2S, pp. 478–480, 2018.

[87] A. Sarkar, A. K. Sahoo, S. Sah, and C. Pradhan, "LSTMSA: A novel approach for stock market prediction using LSTM and sentiment analysis," in *Proc. Int. Conf. Comput. Sci., Eng. Appl. (ICCSEA)*, Mar. 2020, pp. 1–6, doi: 10.1109/ICCSEA49143.2020.9132928.

[88] J. M.-T. Wu, L. Sun, G. Srivastava, and J. C.-W. Lin, "A novel synergetic LSTM-GA stock trading suggestion system in Internet of Things," *Mobile Inf. Syst.*, vol. 2021, pp. 1–15, Jul. 2021, doi: 10.1155/2021/6706345.

[89] C.-R. Ko and H.-T. Chang, "LSTM-based sentiment analysis for stock price forecast," *PeerJ. Comput. Sci.*, vol. 7, pp. 1–23, Mar. 2021, doi: 10.7717/peerj-cs.408.

[90] Y. Qi, Q. Li, H. Karimian, and D. Liu, "A hybrid model for spatiotemporal forecasting of PM2.5 based on graph convolutional neural network and long short-term memory," *Sci. Total Environ.*, vol. 664, pp. 1–10, May 2019, doi: 10.1016/j.scitotenv.2019.01.333.

[91] C.-J. Huang and P.-H. Kuo, "A deep CNN-LSTM model for particulate matter ($PM_{2.5}$) forecasting in smart cities," *Sensors*, vol. 18, no. 7, p. 2220, Jul. 2018, doi: 10.3390/s18072220.

[92] S. Jung, J. Park, and S. Lee, "Polyphonic sound event detection using convolutional bidirectional lstm and synthetic data-based transfer learning," in *Proc. ICASSP IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2019, pp. 885–889, doi: 10.1109/ICASSP.2019.8682909.

[93] W. Zhao, "Novel convolution and LSTM model for forecasting $PM_{2.5}$ concentration," *Int. J. Performability Eng.*, vol. 15, no. 6, pp. 1528–1537, 2019, doi: 10.23940/ijpe.19.06.p4.15281537.

[94] F. Karim, S. Majumdar, H. Darabi, and S. Chen, "LSTM fully convolutional networks for time series classification," *IEEE Access*, vol. 6, pp. 1662–1669, 2018, doi: 10.1109/ACCESS.2017.2779939.

[95] C. Guo, W. Guo, C.-H. Chen, X. Wang, and G. Liu, "The air quality prediction based on a convolutional LSTM network," in *Proc. Web Inf. Syst. Appl., 16th Int. Conf.*, vol. 11817, 2019, pp. 98–109, doi: 10.1007/978-3-030-30952-7_12.

[96] H.-G. Shin, I. Ra, and Y.-H. Choi, "A deep multimodal reinforcement learning system combined with CNN and LSTM for stock trading," in *Proc. Int. Conf. Inf. Commun. Technol. Converg. (ICTC)*, Oct. 2019, pp. 7–11, doi: 10.1109/ICTC46691.2019.8939991.

[97] J. Liu, T. Zhang, Y. Gou, X. Wang, B. Li, and W. Guan, "Convolutional LSTM networks for seawater temperature prediction," in *Proc. IEEE Int. Conf. Signal, Inf. Data Process. (ICSIDP)*, Dec. 2019, pp. 1–5, doi: 10.1109/ICSIDP47821.2019.9173301.

[98] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 1–9, doi: 10.1109/CVPR.2015.7298594.

[99] X. Zhan, Y. Li, R. Li, X. Gu, O. Habimana, and H. Wang, "Stock price prediction using time convolution long short-term memory network," in *Proc. Knowl. Sci., Eng. Manag., 11th Int. Conf.*, vol. 11061, 2018, pp. 461–468, doi: 10.1007/978-3-319-99365-2_41.

[100] D. Li, J. Liu, and Y. Zhao, "Forecasting of $PM_{2.5}$ concentration in Beijing using hybrid deep learning framework based on attention mechanism," *Appl. Sci.*, vol. 12, no. 21, p. 11155, Nov. 2022, doi: 10.3390/app122111155.

TAFIA HASNA PUTRI received the bachelor's degree from the Department of Statistics, Padjadjaran University, where she is currently pursuing the master's (Master of Statistics) degree with the Fast-Track Program. She was a Data Scientist Intern with PT Telkom Indonesia, Telkom, using R programming languages and SQL tools. Her current research interests include statistics is time series data analysis, multivariate data analysis other than that, and the big data field is to study and work in professional data roles, especially machine learning and data science.

REZZY EKO CARAKA (Member, IEEE) is currently an Associate Researcher with the Research Center for Data and Information Sciences, Research Organization for Electronics and Informatics, National Research and Innovation Agency (BRIN), Indonesia, a role he assumed, in February 2022. His prior appointments include positions as a Postdoctoral Researcher with the Department of Statistics, Seoul National University, from December 2019 to December 2021, and a Postdoctoral Researcher with the Department of Nuclear Medicine, Seoul National University Hospital, from January 2021 to January 2022. Following these roles, he assumed the responsibilities of a Research Assistant Professor with the Department of Statistics, Seoul National University, from January 2022 to April 2022. In tandem with his dedicated research pursuits, he has made noteworthy contributions to academia. Since 2021, he has been holding positions as an Adjunct Lecturer with the Faculty of Economics and Business, Universitas Indonesia, and the Graduate School, Department of Statistics, Padjadjaran University. Additionally, he holds the esteemed role of a Senior Research Fellow with the Department of Mathematics, Ulsan National Institute of Science and Technology, South Korea, a position he has maintained, since 2022. Furthermore, he has been contributing as a Visiting Professor with the Department of Big Data Convergence, Pukyong National University, since January 2023. He has also been a Senior Lecturer with the School of Economics and Business, Telkom University, since January 2024. His research interests include statistics, large-scale optimization, machine learning, big data analytics, data science, and sustainable development goals. Acknowledged for his outstanding contributions, he has earned distinction, securing a position in the top 2% of scientists worldwide in A.I., as recognized by Stanford University.

TONI TOHARUDIN received the M.Sc. degree from Katholieke Universiteit Leuven, in 2005, and the Ph.D. degree in spatial sciences from the University of Groningen, in 2010. He is currently a Professor with the Department of Statistics, Universitas Padjadjaran. He acted as the Head of the Research Group in Time Series and Regression. His current research interest includes statistics.

YUNHO KIM is currently an Associate Professor with the Department of Mathematical Sciences, Ulsan National Institute of Science and Technology, Republic of Korea. His research interests include the mathematical understanding of image data, especially medical and biomedical data. The research projects, he either had finished or is still pursuing include image denoising/deblurring/segmentation problems and A.I. research for image processing tasks using reservoir computing networks in neuromorphic computing. The generalized eigenvalue problems and their numerical computations, medical/biomedical image reconstruction problems, and variants of the Allen-Cahn equation in connection with (volume preserving) mean curvature motion.

RUNG-CHING CHEN (Member, IEEE) received the B.S. degree from the Department of Electrical Engineering, National Taiwan University of Science and Technology, Taipei, Taiwan, in 1987, the M.S. degree from the Institute of Computer Engineering, National Taiwan University of Science and Technology, in 1990, and the Ph.D. degree in computer science from the Department of Applied Mathematics, National Chung Hsing University, in 1998. He is currently a Distinguished Professor with the Department of Information Management, Chaoyang University of Technology, Taichung, Taiwan. He is listed in the top 2% of scientists worldwide in A.I. by Stanford University. His research interests include network technology, pattern recognition, knowledge engineering, the Internet of Things, data analysis, and artificial intelligence.

PRANA UGIANA GIO is currently the Founder of STATCAL (statistical software) (https://statcal.com/) and a Content Creator on the Youtube channel: STATKOMAT (programming statistics). He is also a Lecturer with the Department of Mathematics, Universitas Sumatera Utara. His field of study is building web based applications using R & Javascript, probability distribution modeling, Monte Carlo simulation, and Bayesian. He has published dozens of books related to programming and statistics.

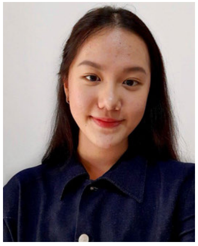
ANJAR DIMARA SAKTI is currently a Lecturer with the Remote Sensing and Geographical Information Sciences Research Group, Bandung Institute of Technology, focusing on the application of remote sensing and spatial data science for developing the geospatial product and global-regional-local policy models to achieve the SDGs concerning the water-food energy-ecosystems nexus.

RESA SEPTIANI PONTOH received the bachelor's degree in statistics from Padjadjaran University and the joint master's degree in business administration and statistical science from Bandung Institute of Technology and La Trobe University. She is currently a Lecturer with the Department of Statistics, Padjadjaran University. Her research interests include econometrics, behavior statistics, and epidemiology.

INDAH RESKI PRATIWI received the bachelor's degree from the Department of Statistics, Padjadjaran University. She was a Laboratory Assistant with the Department of Statistics, Padjadjaran University, in exploratory data analysis, categorical data analysis, multivariate data analysis, and time series data analysis subjects using the R and Python programming languages, from August 2020 to December 2022. She was a Data Analyst Student in the independent study Generasi Gigih 2.0 organized by Yayasan Anak Bangsa Bisa and GoTo Group, from February 2022 to July 2022. She was a Teaching Assistant in a data science class with the Pacmann AI Academy focused on a deep understanding of SQL and Shell Tooling, from May 2022 to July 2022. Her current research interests include deepening her knowledge in data science, machine learning, and big data.

FARID AZHAR LUTFI NUGRAHA received the bachelor's degree from the Department of Statistics, Padjadjaran University. He was an AI and Big Data Research Assistant with the Department of Statistics, Padjadjaran University, conducting research on segmenting 3D images by applying the deep CNN model for segmentation (3DUNet), from August 2022 to December 2022. He was also a Machine Learning Path Student with Bangkit 2022, led by Google, and received the TensorFlow Developer Certificate, from February 2022 to August 2022. He was a Laboratory Assistant with the Department of Statistics, Padjadjaran University, in time series data analysis subject, demonstrating time series forecasting using various NN models with TensorFlow and Python, from August 2022 to December 2022. His current research interests include studying and working in professional data roles, especially machine learning, data science, and big data.

THALITA SAFA AZZAHRA received the bachelor's degree from the Department of Statistics, Padjadjaran University. She was a Machine Learning Path Student with Bangkit 2022, led by Google, Gojek, Tokopedia, and Traveloka, and received the Tensorflow Developer Certificate, from February 2022 to August 2022. She was also a Data Science Intern with PT Erajaya Swasembada, Telkom, who made the dashboard for employee internal assessment, from August 2022 to December 2022. Her current research interests include learn more about data, especially data analysis, data science, machine learning, and big data.



JESSICA JESSLYN CERELIA received the degree (Hons.) in machine learning from Bangkit Academy, in 2022, and the bachelor's degree from the Department of Statistics, Padjadjaran University. She received a Google Certified TensorFlow Developer. She was a Google-led Program in collaboration with GoTo and Traveloka, from February 2022 to August 2022. She was an Assistant Lecturer with the Department of Statistics and Mathematics, Padjadjaran University, in parametric statistics course, from August 2021 to December 2021. Her current research interests include learning more about data analytics, data science, machine learning, and big data.

GUMGUM DARMAWAN received the bachelor's degree in statistics from Padjadjaran University, the master's degree in statistical science from Institut Teknologi Sepuluh November (ITS), Surabaya, and the Ph.D. degree in mathematics from Gadjah Mada University, in 2017. He is currently a Senior lecturer with the Department of Statistics, Padjadjaran University. His research interests include time series analysis, statistics computation, and queueing systems.

DEFI YUSTI FAIDAH received the bachelor's and master's degrees in statistical science from Institut Teknologi Sepuluh Nopember (ITS), Surabaya, where she is currently pursuing the Ph.D. degree in statistics. She is also a Lecturer with the Department of Statistics, Padjadjaran University. Her research interests include climate modeling, extreme events, and ensemble calibration.

BENS PARDAMEAN received the bachelor's degree in computer science and the master's degree in computer education from California State University at Los Angeles, USA, and the Ph.D. degree in informatics research from the University of Southern California (USC). He has over 30 years of global experience in information technology, bioinformatics, and education. His professional experience includes being a Practitioner, a Researcher, a Consultant, an Entrepreneur, and a Lecturer. He currently holds a dual appointment as the Director of the Bioinformatics & Data Science Research Center (BDSRC) | AI Research & Development Center (AIRDC) and a Professor in computer science with Bina Nusantara (BINUS) University, Jakarta, Indonesia.