Multi-sensor weather data prediction-based artificial neural network algorithm

Indra Riyanto [a,b,1,*], Erica Reva Akilah [a,2], Akhmad Musafa [a,3], Eka Purwa Laksana [a,4], Nazori Agni [a,5]

[a] Program Studi Teknik Elektro, Fakultas Teknik, Universitas Budi Luhur, Jakarta 12260, Indonesia
[b] Centre for Remote Sensing Studies, Universitas Budi Luhur, Jakarta12260, Indonesia
[1] indra.riyanto@budiluhur.ac.id; [2] 2152500159@student.budiluhur.ac.id; [3] akhmad.musafa@budiluhur.ac.id;
[4] eka.purwalaksana@budiluhur.ac.id; [5] nazori@budiluhur.ac.id
* corresponding author

ARTICLE INFO

Article history
Received December 17, 2024
Revised October 11, 2025
Accepted October 17, 2025
Available Online May 31, 2026

Keywords
Weather
Numerical Prediction
Classification
Artificial Neural Network
Multi-sensor Data

ABSTRACT

Accurate weather prediction is essential for supporting various human activities and mitigating the impacts of changing atmospheric conditions. Recent advances in artificial intelligence have enabled the development of data-driven forecasting models capable of capturing complex relationships among meteorological variables. This study proposes an Artificial Neural Network (ANN)-based weather prediction model using multi-sensor weather data, including temperature, humidity, precipitation, solar irradiance, and wind velocity. The proposed ANN architecture consists of an input layer, two hidden layers, and an output layer. Model performance was evaluated using three training–testing data splits (90/10, 80/20, and 70/30) with 100 and 150 training epochs. Prediction performance was assessed using accuracy and Root Mean Squared Error (RMSE). Experimental results demonstrate that the proposed model achieves the best performance with a 70/30 training–testing split and 150 training epochs, providing the highest prediction accuracy and the lowest RMSE among the evaluated configurations. These findings indicate that a relatively simple ANN architecture can effectively model multi-sensor weather data and provide reliable weather predictions.

1. Introduction

The city of Jakarta has a wet-and-dry tropical climate with rainy and dry seasons. It also has the potential for extreme weather, such as heavy rain, lightning storms, and strong winds, which can occur and increase the risk of flooding in several areas of the city. Accurate weather monitoring is needed for planning because weather conditions limit the activities that can be carried out; for example, rainfall predictions can help anticipate rain and floods. Weather prediction is a scientific process that estimates the atmospheric state at a given location and time. In weather predictions, many conditions can be observed, such as air temperature, humidity, wind speed and direction, rainfall, and solar irradiation.

Weather prediction methods usually rely on observed event patterns. For example, if the sunset is very red, the next day often brings sunny weather. However, not all of these predictions have been proven reliable [1]; current weather prediction in Indonesia still largely relies on temperature and humidity data processed using basic statistical methods, which often leads to inaccurate weather forecasts [2]. Therefore, an accurate weather prediction system is needed. Predictions can be made in a variety of ways, one of which is using Artificial Intelligence (AI), the development of a computer system capable of performing tasks that require human intelligence [3]. Artificial Neural Networks (ANN) is model inspired by the structure of the human brain, consisting of interconnected nodes [4]. These can then be implemented to improve weather predictions in Indonesia. Our study explores the optimal basic configuration for an ANN to predict weather using air temperature and humidity data patterns.

In several previous studies, the application of the Machine Learning method has been carried out on weather forecasting [5], [6], [7], [8], [9], the design of the weather forecasting system based on Fuzzy Logic [10], and weather forecasting using linear regression using Python [1]. Previous studies have reviewed various methods for rainfall prediction and the problems that may arise when applying different approaches. Nonlinear relationships in rainfall data and the ability to learn from past data make artificial neural networks a preferred approach among all available approaches [11], [12].

Weather is the state of the atmosphere at a given time and in a relatively narrow area over a short period of time [13]. Meteorological data from the atmosphere at any given time, averaged over up to 24 hours, is called weather data. The weather is determined by many factors, including pressure, wind speed, rainfall, temperature, atmospheric phenomena, time, and area [14]. Weather elements are physical factors used to express weather quantities that exhibit unique patterns in each region. This research aims to contribute to weather forecasting in the Jakarta area by using ANN predictions based on 1-year seasonal historical data.

2. Method

Weather stations provide data for various applications, from agriculture to urban planning. However, with the rapid changes in environmental conditions and the increasing complexity of weather patterns, there's a growing need for enhanced measurements and innovative monitoring techniques [15]. Temperature sensors, such as digital thermometers, are used to record air temperatures. The hygrometer measures relative humidity, and the anemometer measures the speed and direction of the wind. A rainfall sensor measures the amount of precipitation falling. Solar irradiation is measured by a pyranometer, and a special UV sensor can measure the level of ultraviolet radiation. Data collected by these sensors were sent to weather stations (Fig. 1), where they were analyzed to make weather predictions and to understand atmospheric conditions. The use of an appropriate, accurate sensor is very important for obtaining reliable data in various meteorological applications and climatological studies.

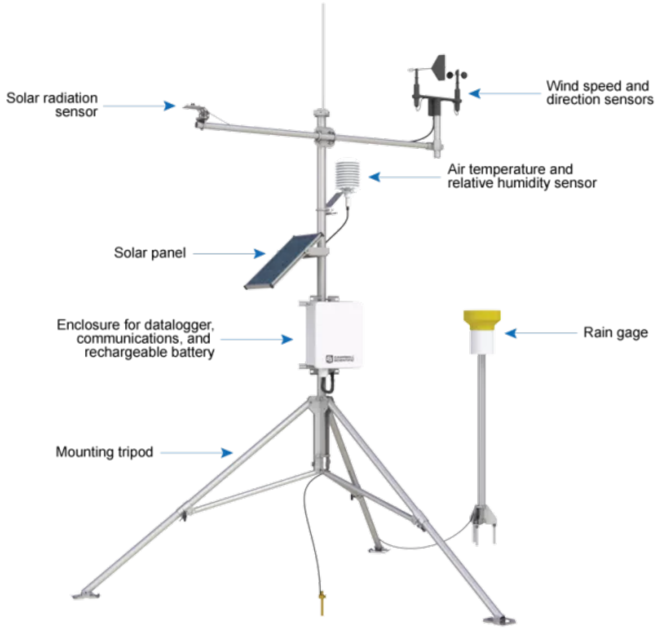


Fig. 1. Example of a Mini Weather Station [16].

2.1. Air Temperature Data

Air temperature is a measure of the warmth or cold of the atmosphere at a given time. Usually expressed in degrees Celsius (°C) or Fahrenheit (°F). Air temperature is measured using a temperature sensor or thermometer. The measured changes are air temperature, differences in the rates of cooling and heating in an area, and the amount of water vapour on the surface of the Earth [17]. According to its place, air temperature varies vertically and horizontally and according to time, every hour, every day, and monthly in a year. The daily variation in air temperatures at the surface of the Earth is mainly controlled by solar radiation; the surface heats up during the day and cools at night [18]. The state of air temperature in a place on the surface of the Earth will be determined by the following factors:

- The length of solar radiation: the longer the sun emits its rays in an area, the more heat is received. A bright atmosphere throughout the day will be hotter than if that day is cloudy from morning.
- The closer the slope of the sun, the more perpendicular the sun's rays are to it; the solar radiation received will be greater, and the temperature in that place will be higher, compared to the place where the sun's position is more tilted.
- The state of clouds in the atmosphere will cause reduced solar radiation received on the surface of the Earth. Because radiation affects the clouds, water vapour in the clouds is absorbed and reflected.

The state of the Earth's surface, differences in land and sea properties, will affect the absorption and reflection of solar radiation. The land surface will absorb and release heat from solar radiation.

2.2. Air Humidity

Air humidity is the concentration of water vapour contained in the air. Water vapour in the atmosphere can change into liquid or solid form, ultimately falling to the earth as rain. This concentration rate can be expressed in terms of absolute humidity, specific humidity, or relative humidity [14]. Relative humidity (RH) is the ratio between the actual amount of water vapour in the air and the maximum amount of water vapour needed to achieve a certain temperature (and pressure). Relative humidity can be stated as follows [8]:

$$R_H = \frac{vapor}{vapor\ capacity} \times 100\% \tag{1}$$

Changes in relative humidity can occur in two main ways: by changing the water vapour in the air and by changing the air's temperature.

2.3. Rainfall

Rainfall is the amount of water in liquid or solid form that falls to the surface of the Earth from the atmosphere within a certain period of time. This is an important component in the hydrological cycle and plays a vital role in ecosystems, agriculture, and human life. Rainfall occurs through a condensation process, in which water vapour in the atmosphere cools and condenses into water droplets or ice crystals. When these drops or crystals join and become quite heavy, they fall to earth as precipitation. This process is influenced by factors such as temperature, air pressure, and humidity.

Types of rainfall clouds can be classified into low liquid water (nimbostratus) and high liquid water (cumulonimbus) clouds, as shown in Fig. 2, and the water will fall to the Earth's surface as:

- Rain: Drops of water that fall when the water vapour in the atmosphere condenses and unites into large drops that are heavy enough to fall to Earth.
- Snow: Ice crystals that fall when the temperature in the atmosphere is very low so that water vapour crystallises into snow.
- Hail: Drops of water that freeze into ice before reaching the surface of the Earth, usually occurring in lightning storms.
- Fog: Very small drops of water suspended in the air near the surface of the earth, reducing visibility.

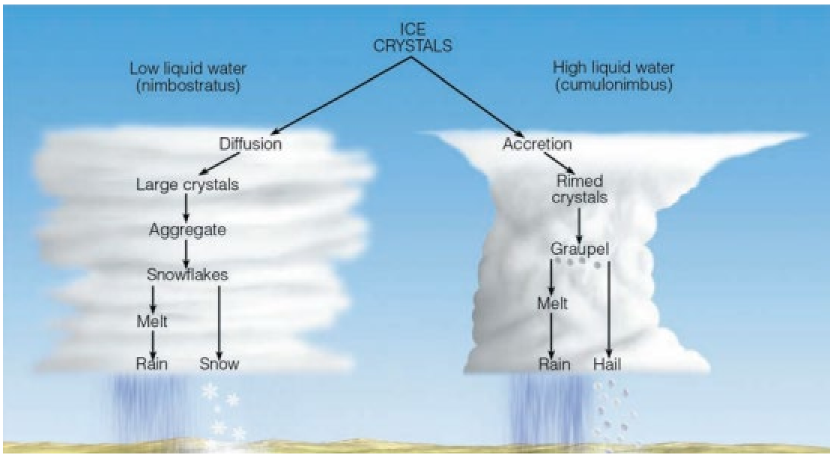


Fig. 2. Rainfall types [19].

Types of rainfall can be classified as follows:

- Rain: Drops of water that falls when the water vapor in the atmosphere condenses and unite into large drops that are heavy enough to fall to earth.
- Snow: Ice crystals that fall when the temperature in the atmosphere is very low so that water vapour crystallises into snow.
- Hail: Drops of water that freeze into ice before reaching the surface of the earth, usually occurs in lightning storms.
- Fog: Very small drops of water suspended in the air near the surface of the earth, reducing visibility.
- Rainfall is measured using a device called a pluviometer or Rain Gauge. This tool collects and measures the amount of precipitation that falls in a certain period of time. The rainfall level measured in unit 1 (one) mm is rainwater as high as 1 (one) mm that falls (accommodated) in a flat area of 1 square meter, with the assumption that nothing evaporates, flows or is absorbed.

Rainfall is typically considered seasonal, meaning that different seasons experience distinct rainfall patterns [20]. Therefore, recognizing rainfall patterns can assist in anticipating natural disasters such as flooding, especially in river basins like Jakarta [21], [22]. Recent developments in weather prediction also include rainfall classification by machine learning [23].

2.4. Solar Irradiation

Solar irradiation is electromagnetic radiation generated by nuclear fusion in the sun's core. This energy includes various types of radiation, such as ultraviolet (UV), visible light, and infrared (IR), for a certain period. The irradiation is expressed in units of power per area. Usually, the value is expressed in watts per square meter (W/m2) [24].

Solar irradiation consists of three primary components: ultraviolet (UV) radiation, visible light, and infrared (IR) radiation. Ultraviolet radiation has wavelengths shorter than visible light and is classified into UVA, UVB, and UVC, with UVC being the most harmful; however, it is largely absorbed by the Earth's atmosphere before reaching the surface. Visible light comprises wavelengths detectable by the human eye, ranging from violet to red, and plays a fundamental role in photosynthesis and the Earth's energy balance. Infrared radiation has wavelengths longer than visible light and is primarily associated with thermal energy, significantly influencing surface temperature, atmospheric heating, and the global climate system. Solar irradiation varies depending on the geographical location, time of day, and season. For example, the area near the equator receives more solar radiation throughout the year than the area closer to the poles [11], [25], [26], [27]. Also, the angle of the sun and the duration of the day affect the intensity of the radiation received.

Radiation entering the Earth's atmosphere undergoes several processes: a portion is scattered by atmospheric particles, is absorbed by atmospheric particles, and is absorbed by the Earth's surface. The total shortwave radiation that reaches the Earth's surface (at the horizontal) is commonly referred to as global or horizontal radiation. This global radiation consists of two types of components, namely the direct radiation component (direct radiation) and the diffuse radiation component (diffuse radiation). In this research, all Sky Surface Shortwave Downward Direct Normal Irradiance (DNI), or irradiation, refers to energy that comes directly from the sun without the spread or refraction of the atmosphere, accounting for atmospheric conditions, including clouds, pollution, and other elements.

2.5. Wind Speed

Wind speed is the rate at which air moves from one place to another. Wind occurs when there is a difference in air pressure, namely maximum air pressure and minimum air pressure. The wind moves from areas of high air pressure to areas of low air pressure. Wind speed is measured using a device called an anemometer. Wind speed is measured as the rate of air movement, usually expressed in units such as kilometres per hour (km/h), meters per second (m/s), or miles per hour (mph) [14]. Wind speed is influenced by several atmospheric and environmental factors, including pressure differences, topography, temperature, and human activities. Air naturally flows from regions of high atmospheric pressure to regions of low pressure, with greater pressure gradients resulting in higher wind speeds. Topographic features such as mountains, valleys, and plains can alter wind patterns by channelling, obstructing, or accelerating airflow in specific areas. Temperature variations between different regions also affect wind speed by creating pressure differences that drive air movement. In addition, human-made structures, including tall buildings and bridges, can modify local wind flow by generating turbulence and altering the wind's direction and velocity..

2.6. Artificial Neural Network (ANN)

Recent trends in weather forecasting have shifted from pure statistics toward machine learning, especially deep learning, as data grow and more parameters are estimated simultaneously [28]. Artificial Neural Networks (ANNs) are a type of artificial neural network that form the basis of deep learning. ANN is a computing system inspired by biological nerve tissue that forms the human brain. ANN can be used to recognize patterns, classify data, and make predictions [29], [30], [31].

Because of the uncertain nature of weather, errors in weather predictions can not be avoided regardless of the modelling approach applied. However, new developments in Machine Learning (ML), especially deep learning (DL), have shown excellent abilities in various prediction processes, such as in medical [32], disaster mitigation [22], [29], and weather applications.

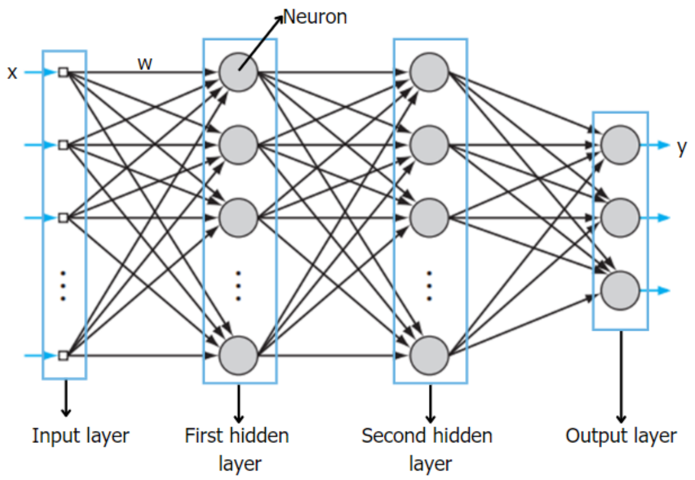


Fig. 3. Artificial Neural Network model [33].

In Fig. 3, circles are called neurons and each neuron runs a function where the number of input values is taken and then mapped into an output value. Arrows illustrate that each neuron's output serves

Vol. 12, No. 2, May 2026, pp. 573-586

as input to the next neuron [33]. In this study, the input consists of measured weather parameters, and the output is future weather parameter values based on these data.

ANNs are generally characterised by an input layer, more than one hidden layer, and one output layer. Fig. 3 shows the conceptual structure of an ANN model used for weather forecasting. The input layer contains input from factors influencing the weather [34], [35], [36]. This information is then processed and analyzed in the hidden layer to determine weights and classification. ANN consists of the following components:

- Neurons (node or perceptron) as the basic unit of ANN. Each neuron receives input, processes it, and produces output. A neuron J receives Input XI with Wij and BIAS BJ weights, then applies the activation function f.
- Layer (layer) consists of the input layer: layers that receive initial input, hidden layers: the layers between the input and output that process data gradually, and the output layer: the layer that produces the final output of the network.
- Weight (weight): each connection between neurons has a weight that determines how strongly one neuron affects other neurons.
- Bias: the value added to the total input of the neuron to adjust the activation function.
- The activation function is a function used to introduce non-linearity into the model, thus allowing the network to study complex relationships. In this paper, the ReLU (Rectified Linear Unit) activation function will be used [3], [4], [9], [29], [37].

ANN can be described in the mathematical model in the following equations:

$$z = \sum_{i=1}^{n} w_i\, x_i + b \tag{2}$$

$$y = f(z) = f\left(\sum_{i=1}^{n} w_i\, x_i + b\right) \tag{3}$$

where $\boldsymbol{x} = [x_1, x_2, \ldots, x_n]$denotes the input vector to the neuron, $\boldsymbol{w} = [w_1, w_2, \ldots, w_n]$ represents the corresponding connection weight vector, $b$is the bias term, and $z$is the weighted sum of the inputs, computed before applying the activation function. The final neuron output is denoted by y, which is obtained by passing $z$through the activation function $f(\cdot)$. In Fig. 3, circles represent neurons, and each neuron performs a function that maps the number of input values into an output value. Arrows illustrate that each neuron's output serves as input to the next neuron [4]. In this study, the input consists of measured weather parameters, and the output is future weather parameter values based on these data.

2.7. System Design

Weather data collected in previous research will then be used to build an ANN model. In ANN modelling, the number of input, hidden, and output layers is determined. At this stage, a simulation of weather prediction programs was designed in Python. Before the ANN model is trained, the data is used to train it to recognize the weather pattern, thereby reducing error. Then, the trained ANN will be tested to determine the accuracy of the developed model. The prediction results obtained will then be analyzed using the data.

2.8. Data Preprocessing

Data preprocessing in Deep Learning refers to the steps taken to prepare and clean the data in Table 1 for model training. The main goal is to improve data quality and ensure the model learns effectively [38].

During the preprocessing stage, several steps were taken to ensure consistency and improve data quality. First, the values in the Hour column were converted into a standard time format to facilitate temporal analysis. The original weather parameter codes obtained from the data source were then replaced with descriptive parameter names, including solar irradiation, temperature, humidity, rainfall, and wind speed, to enhance readability and interpretation. Finally, the solar irradiation measurements, originally expressed in MJ/h, were converted into W/m² by multiplying each value by 277.78, according to Eq. (4) [39]. This conversion standardized the irradiation data into a widely used unit for meteorological analysis and machine learning applications .

$$1\frac{MJ}{h} = \frac{10^6 J}{3600s} = 277.78W \tag{4}$$

Table 1. 12-hour Example of Original Data.

YEAR	MO	DY	HR	ALLSKY_SFC_ SW_DNI (MJ/hr)	T2M	RH2M	PRECTOTCORR	WS2M
2020	1	1	0:00	0	26.93	92.19	4.14	3.21
2020	1	1	1:00	0	26.7	92.94	4.18	3.85
2020	1	1	2:00	0	26.5	93.25	4.22	4.45
2020	1	1	3:00	0	26.37	93.38	4.25	4.61
2020	1	1	4:00	0	26.3	93.19	4.32	4
2020	1	1	5:00	0	26.26	92.88	4.39	3.54
2020	1	1	6:00	0.12	26.37	92.12	4.57	3.34
2020	1	1	7:00	0.09	26.63	91.12	4.79	1.87
2020	1	1	8:00	0.1	26.87	89.5	5.08	1.89
2020	1	1	9:00	0.12	27.09	87.31	5.23	1.78
2020	1	1	10:00	0.22	27.33	85.25	4.95	1.91
2020	1	1	11:00	0.39	27.56	83.38	4.55	1.87
2020	1	1	12:00	0.4	27.67	82.44	4.27	1.78

The preprocessed data is shown in Table 2. After the data has been formatted to meet requirements, the next step in preprocessing is to cleanse it of unnecessary values so the process receives only relevant data. The data that has been cleaned is then normalized using the Min-Max Scaler by calculating the minimum and maximum value of the data while applying the transformation [22].

Table 2. Pre-processed data from Table 1.

Y	M	D	HR	ALLSKY_SFC _SW_DNI (MJ/hr)	Irradiation (W/m^2)	Temperature	Humidity	Rainfall	Wind
2020	1	1	0:00	0	0	26.93	92.19	3.21	4.14
2020	1	1	1:00	0	0	26.7	92.94	3.85	4.18
2020	1	1	2:00	0	0	26.5	93.25	4.45	4.22
2020	1	1	3:00	0	0	26.37	93.38	4.61	4.25
2020	1	1	4:00	0	0	26.3	93.19	4	4.32
2020	1	1	5:00	0	0	26.26	92.88	3.54	4.39
2020	1	1	6:00	0.12	33.3336	26.37	92.12	3.34	4.57
2020	1	1	7:00	0.09	25.0002	26.63	91.12	1.87	4.79
2020	1	1	8:00	0.1	27.778	26.87	89.5	1.89	5.08
2020	1	1	9:00	0.12	33.3336	27.09	87.31	1.78	5.23
2020	1	1	10:00	0.22	61.1116	27.33	85.25	1.91	4.95
2020	1	1	11:00	0.39	108.3342	27.56	83.38	1.87	4.55
2020	1	1	12:00	0.4	111.112	27.67	82.44	1.78	4.27

Min-Max Scaler changes features by changing values in the range of features to a certain scale, usually into the range [0, 1]. This process is done by changing each x feature with equation (5):

$$x_{scaled} = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{5}$$

Next step is to divide the dataset that has been normalized into training and testing data. This process aims to divide the dataset into training and testing sets while preserving the order of the data, which is important for accurate predictions. In this paper, the data will be split using training compositions of 90/10, 80/20, and 70/30.

2.9. Proposed ANN Model

The artificial neural network model consists of input, hidden, and output layers. Input Layer: In this study, the weather-layer parameters process the input, which is then passed to the output layer, which contains the predicted weather. The artificial neural network model is built by determining the input variable, the number of hidden layers, the number of outputs, learning rates, activation functions, momentum (between 0.1 and 0.9), and the number of nodes in each hidden layer. In this study, the ANN model had 5 input and 5 output layers.

The input layer consists of previous weather parameters, and the output layer predicts the upcoming weather parameters. The first step in building the ANN model is to design the network architecture. In this study, 2 hidden layers with 32 and 64 neurons are used [21]. The ANN model consists of 4 layers: 1 input layer, 2 hidden layers, and 1 output layer. In the input layer, there are 5 neurons; in the first hidden layer, there are 32 neurons; in the second hidden layer, there are 64 neurons; and in the output layer, there are 5 neurons (Fig. 4). The data used are hourly weather parameters collected from January 1, 2020 to December 31, 2023.

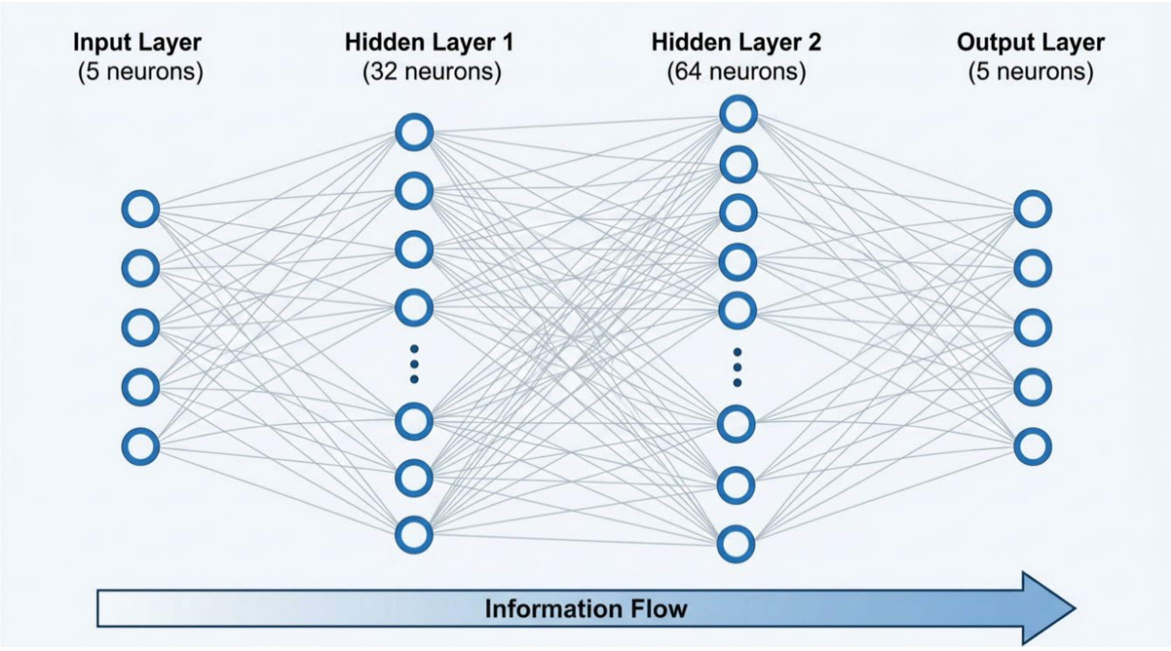


Fig. 4. Proposed ANN Model

3. Results and Discussion

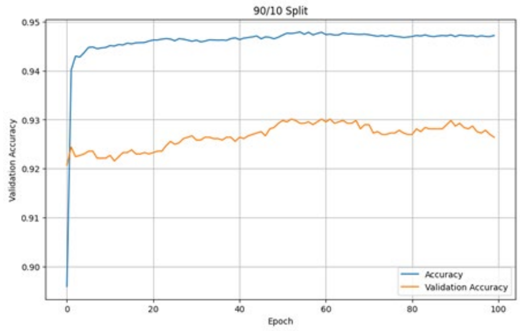
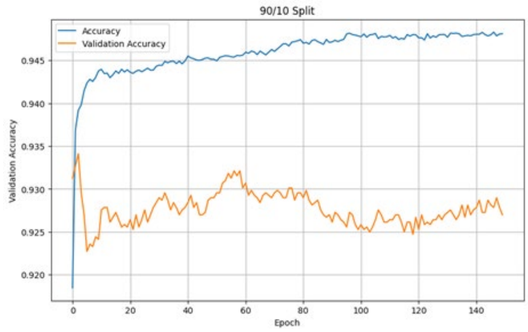
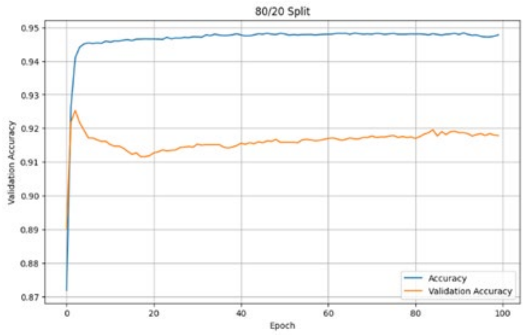
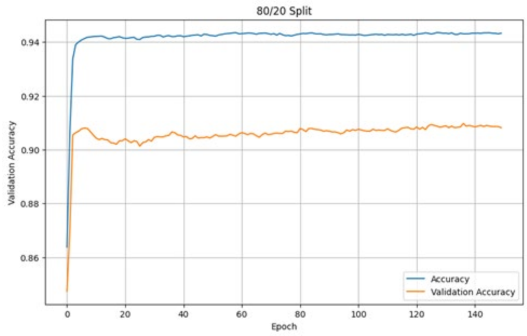
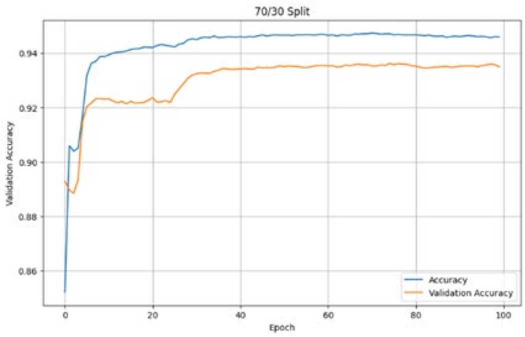
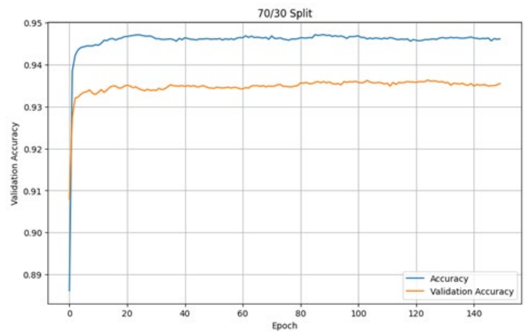
ANN algorithm testing is carried out with three variations in the proportions of training and testing data: 70/30, 80/20, and 90/10 [40]. The amount of data for each composition variation is as in Table 3.

Table 3. Training and Testing Data.

Training/Testing	No. of Data	
	Training	Testing
90/10	31556	3506
80/20	28050	7012
70/30	24543	10519

For each variation in the test data composition, two different epoch values, 100 and 150, are used. The results are displayed in the graphs in Table 4 and Table 5. The evaluation of the ANN algorithm's performance using two different numbers of epochs (100 and 150) and three data-distribution variations (90/10, 80/20, and 70/30).

Table 4. Accuracy testing.

Data Split	Epochs	
	100	**150**
90/ 10		
80/ 20		
70/ 30		

In the first test scenario with 100 epochs, as shown in Table 5, all data distributions (90/10, 80/20, and 70/30) show that training accuracy exceeds 94% and remains stable throughout the 100 epochs. In the 90/10 data distribution, accuracy reaches and remains stable at around 0.9267. Meanwhile, with an 80/20 data split, accuracy decreased to 0.916. In the 70/30 data distribution, the best accuracy at epoch 100 was 0.9301. Root Mean Square Error (RMSE) are calculated at 0.0479 for the 90/10 split, 0.0491 for the 80/20 split, and 0.0455 for the 70/30 split. The training and testing process at 100 epochs was completed in 6 minutes.

Table 5. Performance comparison.

No. of *Epochs*	Metrics	90/10	80/20	70/30
100	Time	6mins	6mins	6mins
	Accuracy	0.9267	0.9161	**0.9301**
	RMSE	0.0479	0.0491	**0.0455**
150	Time	10mins	10mins	10mins
	Accuracy	0.9277	0.9057	**0.9347**
	RMSE	0.0480	0.0577	**0.0419**

In the second test scenario with 150 epochs, across all data distributions (90/10, 80/20, and 70/30), training accuracy quickly reached over 94% (Table 4) and remained stable for the full 150 epochs, as in the previous scenario. In the 90/10 data split (Table 5), there was a slight increase in accuracy to around 0.9277; meanwhile, for the 80/20 split, there was a slight decrease in accuracy to 0.9057, and for the 70/30 split, the accuracy remained the same as in testing with 100 epochs. RMSE for the 90/10 data split was 0.0480; for the 80/20 split, it was 0.0577; and for the 70/30 split, it was 0.0419. The training and testing process at 150 epochs was completed in 10 minutes.

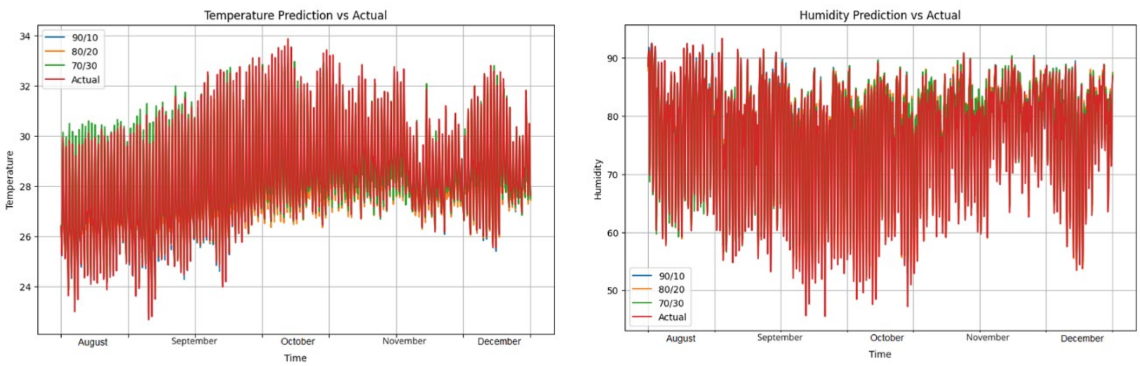
Table 6 shows test results compared with the actual average annual values. It can be seen that the temperature prediction results at epochs 100 and 150 across all data-distribution variations differ little. Testing with 100 or 150 epochs across all data distribution variations yields predictive results close to the actual value, with an error of no more than 2°C.

Table 6. Air temperature prediction accuracy.

No. Of *Epochs*	Metrics	90/10	80/20	70/30
100	Accuracy	0.9884	0.9848	**0.9879**
	RMSE	0.4766	0.5486	**0.4593**
150	Accuracy	0.9892	0.9788	**0.9891**
	RMSE	0.4465	0.7627	**0.4254**

The quantitative performance of the air temperature prediction model is shown in Table 6. All data distribution scenarios achieved high prediction accuracy exceeding 97.8%, with RMSE values below 0.8°C, indicating excellent agreement between the predicted and observed temperatures. At 100 epochs, the 80/20 split achieved the highest accuracy (0.9848), while the 70/30 split produced the lowest RMSE (0.4593°C). After increasing the training to 150 epochs, the 90/10 split attained the highest accuracy (0.9892), whereas the 70/30 split yielded the lowest RMSE (0.4254°C). Overall, although the differences among the three data distributions are relatively small, the 70/30 split consistently demonstrates the lowest prediction error, indicating better generalization performance despite not always achieving the highest numerical accuracy.

Fig. 5 shows the comparison between the actual values and the predicted values for (a) temperature and (b) humidity using three data distribution scenarios (90/10, 80/20, and 70/30). For both parameters, the predictions produced by the three data splits exhibit similar temporal patterns and closely follow the actual observations throughout the evaluation period. However, the 70/30 data split demonstrates the closest agreement with the actual values, particularly in capturing fluctuations and peak variations, indicating its superior predictive performance compared with the 80/20 and 90/10 data distributions..



(a) **(b)**

Fig. 5. (a) Temperature and (b) Humidity prediction vs actual values

The performance of the humidity prediction model under different data distribution scenarios (90/10, 80/20, and 70/30) using 100 and 150 training epochs (Table 7). All models achieved high prediction accuracy (>96%), demonstrating the effectiveness of the proposed approach in modeling humidity variations. At 100 epochs, the 90/10 data split produced the highest accuracy (0.9715), while the 70/30 split achieved a comparable accuracy (0.9712). After increasing the training to 150 epochs, the 90/10 split further improved its performance, reaching the highest accuracy (0.9776) and the lowest RMSE (2.2498). Although the 70/30 split exhibited slightly lower numerical accuracy than the 90/10 split, its prediction curves more closely matched the observed humidity fluctuations, indicating better generalization to unseen data. These results suggest that the proposed model provides reliable humidity predictions across different train–test distributions during the August–December observation period.

Table 7. Humidity prediction.

No. Of *Epoch*s	Metrics	90/10	80/20	70/30
100	Accuracy	0.9715	0.9687	**0.9712**
	RMSE	2.6733	2.9942	**2.7602**
150	Accuracy	0.9776	0.9639	**0.9710**
	RMSE	2.2498	3.5439	**2.8091**

The experimental results demonstrate that the proposed model can accurately predict both air temperature and humidity during the August–December observation period, with prediction trends closely following the actual measurements. Although the evaluation was limited to data collected during Indonesia's wet season, the results indicate the robustness of the proposed approach. Future work should validate the model using full-year and multi-year datasets to assess its performance under a wider range of seasonal and climatic conditions.

4. Conclusion

This study developed an artificial neural network (ANN)-based model to predict air temperature and humidity using different train–test data splits (90/10, 80/20, and 70/30) and training epochs (100 and 150). The experimental results demonstrate that the proposed model achieved high predictive performance for both meteorological parameters across all configurations. Although the 90/10 split produced the highest numerical accuracy in some cases, the 70/30 split consistently yielded lower prediction errors and prediction curves that more closely followed the observed data, indicating better generalization performance. Considering both prediction accuracy and computational efficiency, the 70/30 data split with 100 training epochs is recommended as the optimal configuration, achieving approximately 98% accuracy for air temperature and 97% for humidity while reducing training time. These findings confirm that the proposed ANN model can effectively predict air temperature and humidity during Indonesia's wet season. Future research will extend the evaluation to dry-season and transition-season datasets to assess the model's robustness under varying climatic conditions. Furthermore, performance may be further enhanced through ANN hyperparameter optimization and by integrating the ANN with advanced deep learning architectures, such as Long Short-Term Memory (LSTM) networks, to better capture temporal dependencies in weather data.

Declarations

Author contribution. All authors contributed equally to the main contributor to this paper. All authors read and approved the final paper.
Funding statement. None of the authors have received any funding or grants from any institution or funding body for the research.
Conflict of interest. The authors declare no conflict of interest.
Additional information. No additional information is available for this paper.

References

[1] C. K. Gomathy, U. S. R. Chowdary, V.M. Reddy, and V. Geetha, "Weather forecasting application using Python," *J. Eng. Comput. Archit.*, vol. 12, no. 3, pp. 44–52, 2022, Accessed: May 28, 2026. [Online]. Available: https://www.researchgate.net/publication/360620450_WEATHER_FORECASTING_APPLICATION_USING_PYTHON.

[2] D. Sucahyono and K. Ribudiyanto, *Cuaca dan Iklim Ekstrim di Indonesia*. Penerbit Puslitbang, Badan Meteorologi, Klimatologi, dan Geofisika, 2013. [Online]. Available: https://scholar.google.com/citations?view_op=view_citation&hl=id&user=NJ0GPBMAAAAJ&citation_for_view=NJ0GPBMAAAAJ:9yKSN-GCB0IC.

[3] H. Mauladdawilah, M. Balfaqih, Z. Balfagih, M. del C. Pegalajar, and E. J. Gago, "Deep Feature Selection of Meteorological Variables for LSTM-Based PV Power Forecasting in High-Dimensional Time-Series Data," *Algorithms*, vol. 18, no. 8, p. 496, Aug. 2025, doi: 10.3390/a18080496.

[4] I. Riyanto, M. Rizkinia, R. Arief, and D. Sudiana, "Three-Dimensional Convolutional Neural Network on Multi-Temporal Synthetic Aperture Radar Images for Urban Flood Potential Mapping in Jakarta," *Appl. Sci.*, vol. 12, no. 3, p. 1679, Feb. 2022, doi: 10.3390/app12031679.

[5] M. Holmstrom, D. Liu, and C. Vo, "Machine Learning Applied to Weather Forecasting," 2016, p. 5. Accessed: May 28, 2026. [Online]. Available: https://cs229.stanford.edu/proj2016/report/HolmstromLiuVo-MachineLearningAppliedToWeatherForecasting-report.

[6] A. Gupta, H. K. Mall, and S. Janarthanan., "Rainfall Prediction Using Machine Learning," in *2022 First International Conference on Artificial Intelligence Trends and Pattern Recognition (ICAITPR)*, IEEE, Mar. 2022, pp. 1–5. doi: 10.1109/ICAITPR51569.2022.9844203.

[7] R. Thamilselvan, M. Prabhu, K. T. Selvi, V. Karthik, and T. Bhuvaneshwaran, "An Extensive Review of Rainfall Prediction using Machine Learning and Deep Learning Techniques," in *2022 Second International Conference on Artificial Intelligence and Smart Energy (ICAIS)*, IEEE, Feb. 2022, pp. 198–205. doi: 10.1109/ICAIS53314.2022.9742746.

[8] K. K, V. N C, P. P, and D. Selvapandian, "An Efficient Rainfall Forecasting System using Machine Learning Methods," in *2023 Second International Conference on Electrical, Electronics, Information and Communication Technologies (ICEEICT)*, IEEE, Apr. 2023, pp. 01–06. doi: 10.1109/ICEEICT56924.2023.10157395.

[9] S. Baran *et al.*, "Machine-learning-based probabilistic forecasting of solar irradiance in Chile," *Adv. Stat. Climatol. Meteorol. Oceanogr.*, vol. 11, no. 1, pp. 89–105, Jun. 2025, doi: 10.5194/ascmo-11-89-2025.

[10] Y. Wang *et al.*, "Flood susceptibility mapping in Dingnan County (China) using adaptive neuro-fuzzy inference system with biogeography based optimization and imperialistic competitive algorithm," *J. Environ. Manage.*, vol. 247, no. October, pp. 712–729, Oct. 2019, doi: 10.1016/J.JENVMAN.2019.06.102.

[11] R. Blaga, A. Sabadus, N. Stefu, C. Dughir, M. Paulescu, and V. Badescu, "A current perspective on the accuracy of incoming solar energy forecasting," *Prog. Energy Combust. Sci.*, vol. 70, no. January, pp. 119–144, Jan. 2019, doi: 10.1016/j.pecs.2018.10.003.

[12] H. Bommala, G. B. Kumar, B. S. Babu, M. Sivajyothi, M. V. Pavan Kumar, and S. J. Basha, "A Review on Ensemble Model-Based Rainfall Prediction Techniques," in *2024 2nd International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT)*, IEEE, Jan. 2024, pp. 1474–1478. doi: 10.1109/IDCIoT59759.2024.10467476.

[13] A. M. Siregar, "Klasifikasi Untuk Prediksi Cuaca Menggunakan Esemble Learning," *PETIR*, vol. 13, no. 2, pp. 138–147, Sep. 2020, doi: 10.33322/petir.v13i2.998.

[14] J.-Y. Shin, B. Min, and K. R. Kim, "High-resolution wind speed forecast system coupling numerical weather prediction and machine learning for agricultural studies — a case study from South Korea," *Int. J. Biometeorol.*, vol. 66, no. 7, pp. 1429–1443, Jul. 2022, doi: 10.1007/s00484-022-02287-1.

[15] A. Rivera, P. Ponce, O. Mata, A. Molina, and A. Meier, "Local Weather Station Design and Development for Cost-Effective Environmental Monitoring and Real-Time Data Sharing," *Sensors*, vol. 23, no. 22, p. 9060, Nov. 2023, doi: 10.3390/s23229060.

[16] A. T. Azevedo, R. D. Coelho, E. Carnervisk, A. Maciel, and R. A. Tabile, "Development of a complete automatic weather station for irrigation management via IoT," Feb. 22, 2024. doi: 10.21203/rs.3.rs-3976074/v1.

[17] S. L. Yazid and R. Wulandari, "Design and Construction of a Mini Weather Station Based on Telemetry: English," Journal ELPER-Tech (Electronic Logic Power-Technology). Accessed: May 28, 2026. [Online]. Available: https://jurnal.ugj.ac.id/index.php/ELPER-Tech/article/view/10025.

[18] I. R. Nugraheni *et al.*, "Evaluating surface temperature variabilities and climate extremes in the dieng plateau over three decades," *Indones. Phys. Rev.*, vol. 8, no. 2, pp. 597–615, May 2025, doi: 10.29303/ipr.v8i2.493.

[19] A. R. Islam bahrul, Suranjan Kumar Das,Muhammad Abul Kalam Mallik, "Simulation of a Thunderstorm Event and Its Thermodynamic Features over Mymensingh, Bangladesh Using WRF Model." Accessed: May 28, 2026. [Online]. Available: https://www.researchgate.net/publication/343426127_Simulation_of_a_Thunderstorm_Event_and_Its_Thermodynamic_Features_over_Mymensingh_Bangladesh_Using_WRF_Model.

[20] C. Hao, A. P. Yunus, S. Siva Subramanian, and R. Avtar, "Basin-wide flood depth and exposure mapping from SAR images and machine learning models," *J. Environ. Manage.*, vol. 297, no. November, p. 113367, Nov. 2021, doi: 10.1016/J.JENVMAN.2021.113367.

[21] J.-Y. Lee and J.-S. Kim, "Detecting Areas Vulnerable to Flooding Using Hydrological-Topographic Factors and Logistic Regression," *Appl. Sci.*, vol. 11, no. 12, p. 5652, Jun. 2021, doi: 10.3390/app11125652.

[22] D. Sudiana *et al.*, "Performance Evaluation of 3-D Convolutional Neural Network for Multitemporal Flood Classification Framework With Synthetic Aperture Radar Image Data," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 18, pp. 3198–3207, 2025, doi: 10.1109/JSTARS.2024.3519523.

[23] B. Li *et al.*, "One-kilometer resolution forecasts of hourly precipitation over China using machine learning models," *Atmos. Sci. Lett.*, vol. 26, no. 3, p. e1297, Mar. 2025, doi: 10.1002/asl.1297.

[24] E. Tarigan, "Rooftop pv system policy and implementation study for a household in indonesia," *Int. J. Energy Econ. Policy*, vol. 10, no. 5, pp. 110–115, Aug. 2020, doi: 10.32479/ijeep.9539.

[25] U. Farooq, M. F. Mushtaq, Z. Ullah, M. T. Ejaz, U. Akram, and S. Aslam, "Time Series Analysis of Solar Power Generation Based on Machine Learning for Efficient Monitoring," *Eng. Reports*, vol. 7, no. 2, p. e70023, Feb. 2025, doi: 10.1002/eng2.70023.

[26] A. Alorf and M. U. G. Khan, "Solar Irradiance Forecasting Using Temporal Fusion Transformers," *Int. J. Energy Res.*, vol. 2025, no. 1, p. 3534500, Jan. 2025, doi: 10.1155/er/3534500.

[27] H. N. Noura, Z. Allal, O. Salman, and K. Chahine, "Efficient Machine Learning-Based Prediction of Solar Irradiance Using Multi-Site Data," *Futur. Internet*, vol. 17, no. 8, p. 336, Jul. 2025, doi: 10.3390/fi17080336.

[28] R. T. Gdeeb, "Weather classification using meta-based random forest fusion of transfer learning models," *Int. J. Adv. Intell. Informatics*, vol. 10, no. 2, p. 186, May 2024, doi: 10.26555/ijain.v10i2.1264.

[29] D. Sudiana *et al.*, "A Hybrid Convolutional Neural Network and Random Forest for Burned Area Identification with Optical and Synthetic Aperture Radar (SAR) Data," *Remote Sens.*, vol. 15, no. 3, p. 728, Jan. 2023, doi: 10.3390/rs15030728.

[30] J. R and S. Selvakumari, "Design of a Prediction Model to Predict Students' Performance Using Educational Data Mining and Machine Learning," in *RAiSE-2023*, Basel Switzerland: MDPI, Dec. 2023, p. 25. doi: 10.3390/engproc2023059025.

[31] S. Wu, Y. Zhao, Y. Wang, J. Chen, T. Zang, and H. Chen, "Convolution Feature Inference-Based Semantic Understanding Method for Remote Sensing Images of Mangrove Forests," *Electronics*, vol. 12, no. 4, p. 881, Feb. 2023, doi: 10.3390/electronics12040881.

[32] K. Yothapakdee, Y. Pugtao, S. Charoenkhun, T. Boonnuk, and K. Tamee, "Finding a suitable chest x-ray image size for the process of Machine learning to build a model for predicting Pneumonia," *Int. J. Adv. Intell. Informatics*, vol. 11, no. 1, p. 25, Feb. 2025, doi: 10.26555/ijain.v11i1.1897.

[33] D. Sudiana *et al.*, "Development of Three-Dimensional Convolutional Neural Network for Urban Flood Classification Using Synthetic Aperture Radar Multi-Temporal Image," in *2023 8th Asia-Pacific Conference on Synthetic Aperture Radar (APSAR)*, IEEE, Oct. 2023, pp. 1–6. doi: 10.1109/APSAR58496.2023.10388541.

[34] X. Sun *et al.*, "A data-to-forecast machine learning system for global weather," *Nat. Commun.*, vol. 16, no. 1, p. 6658, Jul. 2025, doi: 10.1038/s41467-025-62024-1.

[35] T. Zhao, Q. Li, T. Tu, and X. Chen, "An extension of WeatherBench 2 to binary hydroclimatic forecasts," *Geosci. Model Dev.*, vol. 18, no. 17, pp. 5781–5799, Sep. 2025, doi: 10.5194/gmd-18-5781-2025.

[36] X. Sun, "A data-to-forecast machine learning system for global weather," Mar. 18, 2025, *Copernicus Meetings*. doi: 10.5194/egusphere-egu25-7911.

[37] Z. Liu, W. Wang, Q. Ma, X. Liu, and J. Jiang, "Rethinking 3D-CNN in Hyperspectral Image Super-Resolution," *Remote Sens.*, vol. 15, no. 10, p. 2574, May 2023, doi: 10.3390/rs15102574.

[38] J. D. Kelleher, *Deep Learning*. The MIT Press, 2019. doi: 10.7551/mitpress/11171.001.0001.

[39] K. AlSharabi, Y. Bin Salamah, M. Aljalal, A. M. Abdurraqeeb, and F. A. Alturki, "Long-Term Forecasting of Solar Irradiation in Riyadh, Saudi Arabia, Using Machine Learning Techniques," *Big Data Cogn. Comput.*, vol. 9, no. 2, p. 21, Jan. 2025, doi: 10.3390/bdcc9020021.

[40] F. Sohil, M. U. Sohali, and J. Shabbir, "An introduction to statistical learning with applications in R," *Stat. Theory Relat. Fields*, vol. 6, no. 1, pp. 87–87, Jan. 2022, doi: 10.1080/24754269.2021.1980261.