

Deep Learning-Based Weather Prediction Model Using BMKG Big Data for Rice Planting Season Optimization in Indonesia

Maradona Jonas Simanullang¹⁾, Elsy Sabrina Asmita Simorangkir²⁾, Herry Daniel Marpaung³⁾, Yudisa Halawa⁴⁾, Tria Adelia Putri Br Gurusinga⁵⁾, Frans Mikael Sinaga^{6*)}

^{1,2,3,4,5)}Faculty of Technology and Health Sciences, Universitas Senior Medan, Indonesia

⁶⁾Informatics Department, Faculty of AI and Data Sciences, Universitas Pelita Harapan, Indonesia

¹⁾maradonajonassimanullang@gmail.com, ²⁾elsyasabrinaas@gmail.com, ³⁾marpaungherry1@gmail.com,

⁴⁾yudisahalawayudisahalawa@gmail.com, ⁵⁾triaadelia97@gmail.com, ⁶⁾frans.sinaga@uph.edu

Submitted : Aug 4, 2026 | Accepted : Aug 24, 2026 | Published : Oct 5, 2026

Abstract: This study develops and evaluates a deep learning-based weather prediction model built on BMKG (Indonesian Meteorological, Climatological, and Geophysical Agency) data, with the aim of informing rice planting season optimization in Deli Serdang Regency, North Sumatra, Indonesia. Two recurrent architectures, Long Short-Term Memory (LSTM) and Bidirectional Long Short-Term Memory (BiLSTM), were compared for daily rainfall prediction using 788 observations collected between January 2024 and March 2026, split chronologically into training, validation, and testing subsets. The LSTM model achieved a Root Mean Square Error (RMSE) of 24.42 mm, a Mean Absolute Error (MAE) of 15.20 mm, and an R^2 of 0.0420, while the BiLSTM model achieved a lower RMSE of 23.64 mm, a lower MAE of 14.89 mm, and a lower R^2 of 0.0233. Although BiLSTM produced marginally smaller errors, LSTM explained a slightly larger share of rainfall variability, and both R^2 values remained close to zero, indicating that neither model captured the full variability of the rainfall series. Because of its higher R^2 , the LSTM-based forecasts were translated into an indicative cropping-calendar classification of wet, normal, and dry months as a preliminary basis for advancing, maintaining, or delaying the onset of the rice planting season in Deli Serdang Regency. Given the limited explanatory power of both models and the absence of baseline forecasts, repeated training, and time-series cross-validation in this study, these planting-season indications should be read as exploratory and complementary to, rather than a replacement for, the operational cropping calendar issued by BMKG and the Ministry of Agriculture. Future research should incorporate longer historical records, additional meteorological predictors, baseline models, and rigorous validation to develop more reliable, data-driven rice planting season decision support for Indonesia.

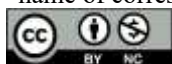
Keywords: Rainfall Prediction, Deep Learning, Long Short-Term Memory (LSTM), Bidirectional Long Short-Term Memory (BiLSTM), BMKG Meteorological Data, Time Series Forecasting, Rice Planting Season Optimization.

INTRODUCTION

Indonesia is a country where farming is a major part of life, and the agriculture sector is very important for making sure the nation has enough food. Rice is one of the most important agricultural products because it is the main food source for most people in Indonesia. Rice production depends a lot on the environment, especially on weather and climate conditions. Rainfall is a key factor that influences how much water is available, how crops grow, and when the best time is to plant them. Changes in how much rain falls because of global climate change have made it harder to predict the seasons, made extreme weather events worse, and increased the chance that crops will not grow as well. Because of this, being able to predict rainfall accurately is very important for helping plan farming activities and lowering the chance of crops failing (Kaplun et al., 2024; Chang, 2024).

In recent years, climate change has caused weather patterns to become more complicated and harder to forecast. Things like heavier rainfalls, longer dry spells, and changes in when to plant crops are causing big problems for

*name of corresponding author



This is anCreative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

farming. Wrong weather information can cause people to use the wrong farming methods, waste water by not managing it properly, and plant crops at the wrong times. So, it's really important to have a weather prediction system that can give accurate and flexible forecasts, especially as the climate keeps changing.

The Meteorological, Climatological, and Geophysical Agency (BMKG) has a large network of weather observation stations throughout Indonesia, always gathering weather data from many monitoring points. In this study, a part of the BMKG weather data from Deli Serdang Regency was used. This part includes 788 daily rainfall records taken from January 2024 to March 2026. The dataset was studied with the help of artificial intelligence and deep learning methods to find out how rainfall changes over time and to help with weather predictions for farming plans.

Quick progress in Artificial Intelligence, especially deep learning, has greatly helped in analyzing time-series data and improving weather forecasts. One of the most commonly used deep learning models is Long Short-Term Memory, or LSTM. It is an improved version of the Recurrent Neural Network, which helps solve the vanishing gradient issue that often happens when learning from sequences of data. LSTM can remember important information from the past for a long time because of its memory cell structure, which helps the model understand and track long-term patterns in rainfall data. Because of these abilities, LSTM has become one of the most commonly used methods in weather prediction and water flow modeling (Necesito et al., 2023; Wang et al., 2023).

Although LSTM works well for understanding long sequences, it only reads information from start to finish, which means it can't use information from later parts of the sequence to better understand the current part. This one-way learning process might limit its ability to handle complicated two-way time connections that exist in weather data over time. To solve this problem, researchers used Bidirectional Long Short-Term Memory (BiLSTM), which looks at information in both forward and backward directions. By including both past and historical information in the input during training, BiLSTM can better understand time-based patterns and may lead to more accurate predictions, especially for data that changes a lot over time (Peeriga et al., 2024; Li et al., 2024).

Many research papers have shown that deep learning methods work well for predicting rainfall. Tanvir and Naiema in 2025 used a model based on LSTM to predict rainfall and found that their forecasts had fairly small errors, showing that this method works well for weather-related uses. Similarly, Necesito et al. (2023) found that LSTM works well for predicting daily rainfall because it can learn long-term patterns over time, especially when there isn't much observational data available. In another study, Wang and their team in 2023 used a framework based on LSTM to estimate precipitation and showed that the model works well in capturing how meteorological variables change over time.

Besides looking at LSTM on its own, many studies have looked at different deep learning models to find out which one works best for making predictions. A study used the NASA POWER weather data to compare the performance of BiLSTM, CNN-BiLSTM, LSTM, and XGBoost models. The results showed that BiLSTM worked better than other models that look at sequences, because it uses information from both past and future time steps, which helps understand how weather patterns change more fully. The hybrid CNN-BiLSTM structure also saw better results because it first captured spatial and local details, then used them to understand time-based patterns, which helped it better detect complicated weather changes.

Peeriga et al. (2024) also suggested a rainfall prediction model that uses BiLSTM, working inside an AIoT system for smart agriculture. Their results showed that BiLSTM was better at capturing complex patterns over time compared to regular one-way models. Jumadi and others in 2025 created an AI method that uses several models together to predict rainfall over time and space. They found that using multiple models together gave better results than using just one model alone.

Even though many past studies have shown that both LSTM and BiLSTM work well for weather prediction, there are still some areas that need more research. How well forecasts work depends a lot on things like the type of data used, how often observations are made, which variables are included, and the weather conditions in a particular area. So, models that work well in one part of the world might not be as accurate in places with different weather conditions. Most of the current research has mainly been about improving the structure of models or making predictions more accurate. However, not many studies have done a full comparison between LSTM and BiLSTM models using local BMKG weather data that directly helps with farming plans, especially for predicting rainfall in Indonesia's weather conditions.

This study looks at using LSTM and BiLSTM models to predict daily rainfall based on local BMKG weather data, with the ultimate aim of supporting rice planting season optimization in Deli Serdang Regency. The study uses 788 rainfall measurements taken from January 2024 to March 2026 and checks how well the predictions work by using RMSE, MAE, MSE, and R squared. The better-performing forecast is then translated into an indicative wet/normal/dry cropping-calendar classification that points toward advancing, maintaining, or delaying the planting season; the study does not build a full crop-water simulation model, a formal decision-optimization algorithm, or a validated agronomic threshold, so the resulting classification should be treated as an exploratory, indicative signal rather than an operational planting recommendation.

This study looks at how well LSTM and BiLSTM models perform on the BMKG rainfall data, not just by comparing their prediction accuracy. The results are only based on the time period that was tested because the model wasn't retrained or checked using time-series cross-validation. The results are shown as a comparison within the same study instead of as proof that the model is generally strong or consistent.

LITERATURE REVIEW

Rainfall Forecasting

Rainfall is a key weather factor that affects many areas, especially farming, water studies, and how we manage water resources. In farming, having accurate information about rainfall is very important for deciding when to plant crops, figuring out how much water is needed for irrigation, and coming up with ways to reduce the chances of losing crops. However, climate change has changed how rain falls and when seasons happen, leading to more unpredictable weather. These changes show that there is an increasing need for better and more dependable methods to predict rainfall, which helps in planning for agriculture and adapting to climate changes (Kaplun et al., 2024; Chang, 2024).

Rainfall forecasting is the process of predicting how much rain will fall in the future by looking at past weather data. Rainfall records are a type of time-based data, which means they show patterns over time. These records have complicated features, such as how rain amounts change over time, how they vary by season, and how different weather factors are connected in non-linear ways. These features make predicting rainfall a tough job and often stop traditional statistical forecasting from working well, since these methods usually have a hard time understanding complex time patterns and nonlinear weather changes (Necesito et al., 2023).

Recent improvements in Artificial Intelligence (AI) have brought more advanced methods for predicting rainfall. These methods allow computer models to study and understand complicated patterns from past rainfall data over time. Deep learning has become very popular because it can automatically find important patterns in raw data and handle complex, non-linear, and sequential relationships effectively. Because of this, deep learning methods are now widely used in weather forecasting and have shown great success in making rainfall predictions more accurate (Wang et al., 2023).

BMKG Meteorological Data

Meteorological observations are a basic source of information used to create weather forecasting systems. In Indonesia, the Meteorological, Climatological, and Geophysical Agency (BMKG) regularly gathers a variety of weather data, such as rainfall, air temperature, humidity, atmospheric pressure, and other related weather factors, using a large number of observation stations. In this study, BMKG meteorological data were used as a local time-series dataset to forecast rainfall in Deli Serdang Regency. The dataset includes 788 daily rainfall measurements taken from January 2024 up to March 2026. The data was looked at using deep learning methods to find patterns over time and check how well the LSTM and BiLSTM models can predict future outcomes. This research is focused on predicting rainfall patterns using meteorological data from BMKG, rather than on handling large amounts of data or analyzing data across multiple systems.

The growing access to longer records of past weather data has opened up new possibilities for creating smart forecasting tools that can find useful patterns in historical weather information. Unlike regular ways of analyzing things, using artificial intelligence helps in understanding complex connections and time-related patterns found in weather information. Because of this, combining BMKG weather data with AI methods has greatly enhanced the ability to predict the weather and offers helpful guidance for managing agriculture, like planning farming activities and assessing risks related to climate (Jumadi et al., 2025).

Deep Learning for Weather Forecasting

Deep learning is a more advanced part of machine learning that uses multiple layers of artificial neural networks to learn complicated patterns directly from data. Unlike traditional machine learning and statistical approaches, deep learning models can automatically find and use important features without needing a lot of manual work to create them. This ability lets them handle very complex relationships and detailed data patterns, which makes them great for tackling tough prediction tasks.

In weather forecasting, deep learning is now widely used because weather data often changes over time, has patterns that repeat every year, and follows complex, unpredictable trends that are hard to model with older forecasting methods. By studying past patterns in data, deep learning models can enhance their ability to predict future events and more accurately capture how weather conditions change over time.

Several types of deep learning models have worked well for weather forecasting tasks. These include Recurrent Neural Networks (RNNs), which are made to work with data that comes in a sequence. There are also Long Short-Term Memory (LSTM) networks, which help in understanding patterns that happen over time. Bidirectional Long Short-Term Memory (BiLSTM) networks use information from both the beginning and the end of a sequence to better understand the context. Another type is a hybrid model like CNN-BiLSTM, which mixes convolutional

neural networks with recurrent networks to improve how features are extracted before they are used for time-based learning (Wang et al., 2023; Li et al., 2024). These methods have shown a lot of promise for better predicting rainfall in tough weather conditions.

Long Short-Term Memory (LSTM)

Long Short-Term Memory, or LSTM, is a kind of Recurrent Neural Network designed to fix a problem called vanishing gradient. This problem occurs when the network has difficulty learning from data that comes in order. Unlike normal RNNs, LSTM has a special memory system inside that helps the network remember important information for a longer time. This ability works because of a memory cell that is managed by three different gates: the forget gate, the input gate, and the output gate. These gates work together to carefully control how information is handled during the learning process. These gates help the network focus on important information and remember the key time-related details that are necessary for making accurate predictions (Tanvir & Naiema, 2025).

LSTM is widely used in deep learning for predicting data that changes over time because it can keep track of information for a long time, which helps with tasks like predicting the weather, estimating how much rain will fall, and keeping an eye on the environment. By looking at how things change over time using past data, LSTM can figure out complex patterns in sequences and give reliable predictions. Previous studies have found that LSTM is effective at making accurate predictions with fewer errors in various weather-related applications, which makes it a good option for predicting rain and other climate-related forecasts (Necesito et al., 2023; Wang et al., 2023).

Even though these benefits are there, LSTM processes sequential data in just one way, typically from the beginning to the end. Because of this, the model cannot use information from observations that take place outside the historical input window during training. This one-way learning approach might not work as well when dealing with two-way time connections that commonly appear in weather data over time, which could prevent it from accurately predicting difficult weather conditions (Peeriga et al., 2024).

Bidirectional Long Short-Term Memory (BiLSTM)

Bidirectional Long Short-Term Memory, or BiLSTM, is an improved version of the standard LSTM network. It helps in understanding sequences by learning information from both the forward and backward directions. Unlike the usual LSTM structure, which only looks at data from the beginning to the end, BiLSTM examines the sequence in both the forward and backward directions. This way of learning allows the model to use information from both past data and events that happen outside the normal time period it's trained on, which helps it better understand longer time connections and more complex patterns in sequences (Li et al., 2024).

In weather forecasting, this skill is very useful because how the atmosphere behaves includes complicated connections that happen over time, not just simple things that directly cause other things. By looking at information from both before and after, BiLSTM can better understand how weather conditions change over time, which helps make weather predictions more accurate. Research done by Peeriga and their team in 2024 found that BiLSTM works better than regular one-way recurrent models for predicting rain. This means BiLSTM is good at understanding complex weather patterns and helps make rain forecasts more accurate.

Comparison of LSTM and BiLSTM

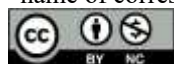
LSTM and BiLSTM are both designed to work with data that changes over time. They both understand how things change by looking at what happened earlier. Both designs use recurrent neural networks and have the same main goal, but they process sequential data in completely different ways. A regular LSTM looks at data in one direction, using only what happened before to guess what will happen next. Unlike other methods, BiLSTM processes sequences in both forward and backward directions at the same time. This allows the model to consider information from the beginning and the end of the sequence, which helps it understand the entire context better. This two-way learning process allows the BiLSTM model to better understand how things are connected over time compared to the regular LSTM method.

Some earlier research suggests that BiLSTM works better when it's used with data sets that change in complicated ways over time. Studies compared different ways to predict outcomes using NASA POWER weather data. The methods looked at were BiLSTM, CNN-BiLSTM, LSTM, and XGBoost. The results show that models using BiLSTM do a better job than regular LSTM at recognizing changes in weather patterns because they can understand connections from both the past and the future more effectively. Adding convolutional layers to the CNN-BiLSTM model helps the network pick up better features first, which makes it easier for the model to learn patterns over time. This results in the model being more skilled at identifying complicated weather changes and leads to more accurate predictions.

Model Evaluation

Model evaluation is an important part of building a predictive model because it helps to see how accurately the model can forecast real-world results. To see how good a model is at making predictions, researchers use several

*name of corresponding author



This is an Creative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

numbers to evaluate its performance. These numbers include Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Square Error (RMSE), and the coefficient of determination (R^2). These metrics provide additional details on how correct and trustworthy the model is.

Among these methods, MAE, MSE, and RMSE are often used to check how incorrect predictions are by comparing the predicted values with the real observed values. In contrast, the coefficient of determination, also called R squared, shows how well a model explains the changes in the outcome variable. A forecasting model is generally considered better if it has lower numbers for MAE, MSE, and RMSE, and if the R^2 value is near 1, which means it can better show the true patterns in the data (Tanvir & Naiema, 2025).

$$MAE = \frac{1}{n} \sum |y - \hat{y}|$$

a. Mean Squared Error (MSE)

Mean Squared Error (MSE) is a method used to show how much the predicted values differ from the real values on average. It does this by first finding the difference between each predicted and actual value, then squaring those differences, and finally averaging all the squared differences. By squaring each error from the prediction before averaging them, MSE treats large mistakes more heavily than small ones. This metric is really useful when evaluating forecasting models, especially when you want to avoid big prediction errors, because it highlights how much unusual data points and large mistakes can impact the model's overall performance. A lower MSE indicates that the predictions are nearer to the actual data, which means the forecast is more accurate.

$$MSE = \frac{1}{n} \sum (y - \hat{y})^2$$

b. Root Mean Square Error (RMSE)

Root Mean Square Error (RMSE) is calculated by taking the square root of Mean Squared Error (MSE). This measure is commonly used to assess how accurate predictions are, as it expresses the error using the same units as the original data. This feature makes RMSE more straightforward to understand in relation to MSE, while still highlighting significant prediction errors. Consequently, a lower RMSE value signifies that the predictions made by the model are more accurate and better correspond to the actual data, indicating that the model is efficient in making forecasts.

$$RMSE = \sqrt{MSE}$$

c. Coefficient of Determination (R^2)

The R squared statistic indicates the extent to which a model can account for the variations in the real data it aims to forecast. It reflects the degree to which the model clarifies the changes in the dependent variable. When the R^2 statistic approaches 1, it signifies that the model accounts for a significant portion of the fluctuations observed in the data, demonstrating its robust predictive capability and overall effectiveness.

METHOD

This research utilizes precipitation data gathered by the Indonesian Meteorological, Climatological, and Geophysical Agency (BMKG) to develop and evaluate deep learning models for forecasting daily rainfall, with a view toward rice planting season optimization. The dataset consists of 788 records and is distributed over time with 70% allocated for training, 15% for validation, and 15% for testing purposes. In addition to rainfall prediction, the methodology includes a simple rule-based step that translates the better-performing model's forecasts into a wet/normal/dry cropping-calendar classification; however, this research does not incorporate a crop growth simulation model, a crop-water requirement calculation, or a formal decision-optimization algorithm.

During the preparation phase, the dataset underwent processes such as addressing missing values, managing outliers, and normalizing data prior to constructing sequences. The current manuscript lacks adequate information to replicate the specific approach taken for handling the BMKG missing-value codes, the method of imputation used, the standards for identifying outliers, or the procedure for fitting the scaler. As a result, these aspects are recognized as limitations for reproducibility and need to be clearly documented in a subsequent revision or additional supporting material.

Two recurrent deep learning architectures, Long Short-Term Memory (LSTM) and Bidirectional Long Short-Term Memory (BiLSTM), are implemented to model the temporal characteristics of the rainfall time series. While the LSTM model learns sequential information in the forward direction, the BiLSTM architecture extends this capability by processing the sequence in both forward and backward directions. This bidirectional learning mechanism enables the model to capture richer contextual information and has been shown to improve forecasting performance for complex temporal datasets (Muhammad Waqas, Humphries, Hlaing, & Ahmad, 2024).

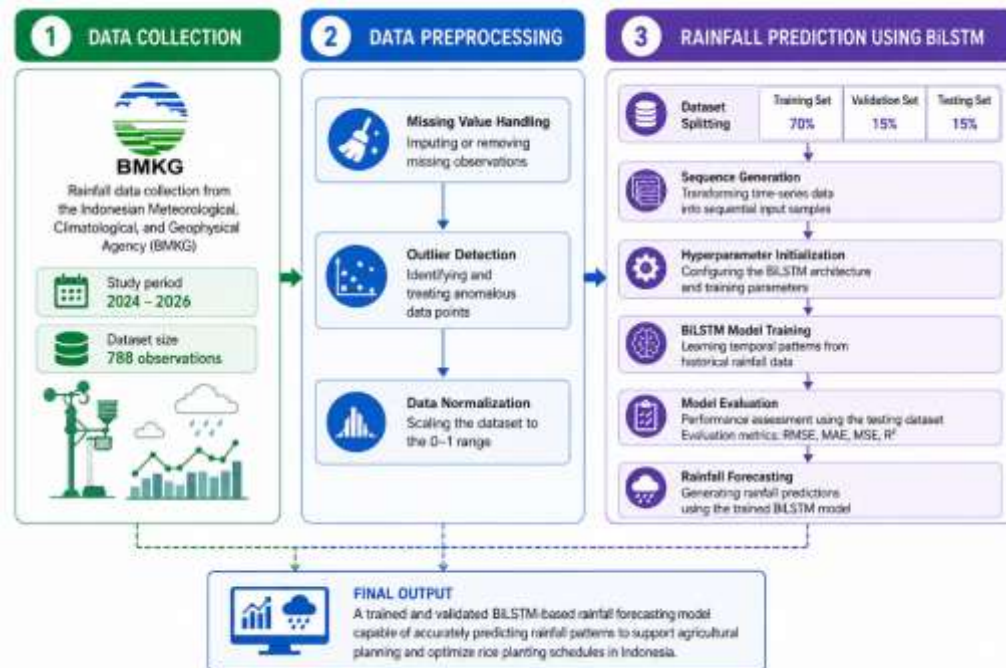


Figure 1. Problem-Solving Approach

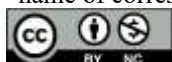
Data Collection

Meteorological information regarding Deli Serdang Regency for the years 2024 to 2026 was sourced from the Indonesian Meteorological, Climatological, and Geophysical Agency, commonly known as BMKG. The choice of Deli Serdang Regency as the focus of this research was due to its status as a key rice-producing area in North Sumatra, Indonesia, highlighting the critical need for precise rainfall predictions to support agricultural planning and rice farming. A portion of the gathered meteorological information can be found in Table 1.

Table 1. Meteorological Data for Deli Serdang Regency

NU	DATE	TN	TX	TAVG	RH_AVG	RR	SS	FF_X	DDD_X	FF_AVG	DDD_CAR
1	06-01-2024	25	31.4	28.7	82		1.7	4	315	2	W
2	07-01-2024	24.2	29.6	27.6	84	7.5	2.7	3	270	1	C
3	08-01-2024	24.6	27.8	25.5	94	0	0	3	180	1	C
4	09-01-2024	24.2	30.8	28.1	83	36	0	4	360	2	C
5	10-01-2024	24.6	31.4	28.1	84	3.4	2.2	4	270	2	C
6	11-01-2024	25.2	32.6	29.2	80	16.7	5.3	5	360	2	N
7	12-01-2024	25.4	32.6	29.8	76		4.3	4	45	2	C
8	13-01-2024	24.2	32.6	29.7	74	0	4.1	4	45	2	NE
9	14-01-2024	25	33	29.6	75		6.9	6	45	2	NE
10	15-01-2024	24.2	32.4	29.1	75		7.6	4	45	2	NE
11	16-01-2024	25	32.2	28.6	81		3	3	45	1	NE
12	17-01-2024	24.8	30.8	27.7	84	5.4	1.7	4	45	1	C
13	18-01-2024	25	29.4	27	89	5	0.5	3	270	1	C
14	19-01-2024	25	29.6	27.3	87	12	0	4	45	1	C
15	20-01-2024	23.8	33	29.6	75	1.7	0.5	3	45	2	NE
16	21-01-2024	25.2	33.2	30.3	72	0	8.1	7	45	3	NE
17	22-01-2024	23.8	33.8	29.7	70		8.7	4	360	2	C
18	23-01-2024	24.2	30.8	28.7	76		7.2	4	45	2	NE
19	24-01-2024	24.4	32.2	29.2	77		3.6	7	45	2	NE
20	25-01-2024	24.2	32.6	29.3	74	1.3	5.3	6	90	2	E

*name of corresponding author



This is anCreative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

21	26-01-2024	24.6	32.8	29.6	74	4.2	5.5	6	45	2	NE
22	27-01-2024	25	32.4	29.5	76	1	7.5	7	90	3	E
23	28-01-2024	24.8	31.8	28.4	82	8888	6.3	3	180	1	C

Data Preprocessing

In the initial phase of data preparation, the dataset was organized to enhance its quality prior to the development of the model. This stage involved addressing absent values, identifying and managing outliers, and implementing data normalization to guarantee that all input variables were aligned on a similar scale. These steps in preprocessing are crucial for facilitating model convergence and boosting the predictive effectiveness of the BiLSTM model.

Dataset Partitioning

Following the preprocessing phase, the organized dataset was split into different subsets intended for model creation and evaluation of its performance. The information was categorized into training, validation, and testing datasets. The training dataset served as the basis for understanding the fundamental rainfall patterns, the validation dataset was utilized to fine-tune model parameters and minimize the chances of overfitting, whereas the testing dataset was set aside to assess the prediction capabilities of the models on data that had not been encountered before.

Sequence Generation

The series of rainfall data as converted into consistent-length historical sequences utilizing a sliding-window methodology. The specific length of the window and the duration for forecasting are not disclosed in the provided experimental details, which hinders reproducibility. In a proper forecasting framework, each sequence aimed at a specific target should include only data points that come before the target; any data points following the start of the forecast should be excluded from both the forward and backward passes.

1. Hyperparameter Initialization

The document indicates that various hyperparameter configurations were assessed, such as learning rate, batch size, the number of hidden units, the choice of optimizer, and the number of training epochs. Nevertheless, the chosen final values, dropout settings, early-stopping guidelines, random seed, exploration range, and tuning methods are not disclosed. As a result, the precise experiment cannot be replicated based solely on the existing manuscript. These configuration specifics need to be included in an upcoming full revision or additional material.

2. BiLSTM Modeling

Bidirectional Long Short-Term Memory (BiLSTM) processes each input sequence in forward and backward directions. This architecture can represent contextual relationships among observations inside the specified input window. However, bidirectional processing does not justify access to observations that occur after the forecast origin. For a forecasting application, the backward pass must operate only within the historical input sequence; otherwise, information leakage may occur. Because the current manuscript does not report the exact sequence construction, the absence of leakage cannot be independently verified from the reported methodology.

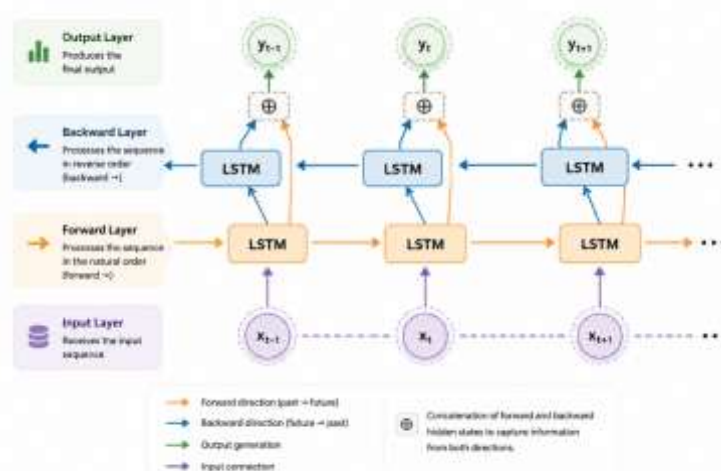
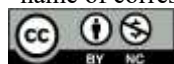


Figure 2. Bidirectional Long Short-Term Memory (BiLSTM) Model Architecture

*name of corresponding author



This is anCreative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

The BiLSTM framework carries out forecasting using three key computational steps: propagation in the forward direction, propagation in the backward direction, and the computation of outputs. These phases allow the model to acquire temporal relationships by handling sequential data in both forward and reverse directions prior to producing the ultimate prediction.

a. *Forward Hidden Layer Computation*

$$\vec{h}_t = f(W_1 x_t + W_2 \vec{h}_{t-1})$$

Where:

$\vec{h}_t$ = forward hidden state at time t

f = activation function

b. *Backward Hidden Layer Computation*

$$\overleftarrow{h}_t = f(W_3 x_t + W_5 \overleftarrow{h}_{t+1})$$

Where:

$\overleftarrow{h}_t$ = backward hidden state at time t

f = activation function

c. *Output Layer Computation*

$$O_t = g(W_4 \vec{h}_t + W_6 \overleftarrow{h}_t)$$

Where:

O_t = output at time t

g = activation function

W = weight parameters

x_t = input data at time t

t = time step

RESULT

This segment outlines the experimental findings for LSTM and BiLSTM models based on 788 rainfall observations from BMKG, collected from January 2024 to March 2026. The evaluation of predictive accuracy was conducted utilizing RMSE, MAE, MSE, and R^2 metrics. The analysis excluded models such as naive persistence, seasonal-naive, historical-mean, ARIMA/SARIMA, Random Forest, or XGBoost as benchmarks. Thus, the results presented solely facilitate a comparison between the two recurrent neural network architectures and do not indicate that deep learning has benefits over more straightforward forecasting methods.

To ensure a fair and consistent comparison, the dataset was partitioned chronologically into 70% for model training, 15% for validation, and 15% for testing. This chronological data splitting strategy preserves the temporal structure of the time-series data while providing an unbiased evaluation of each model's forecasting capability on previously unseen observations.

Performance on the BMKG Dataset

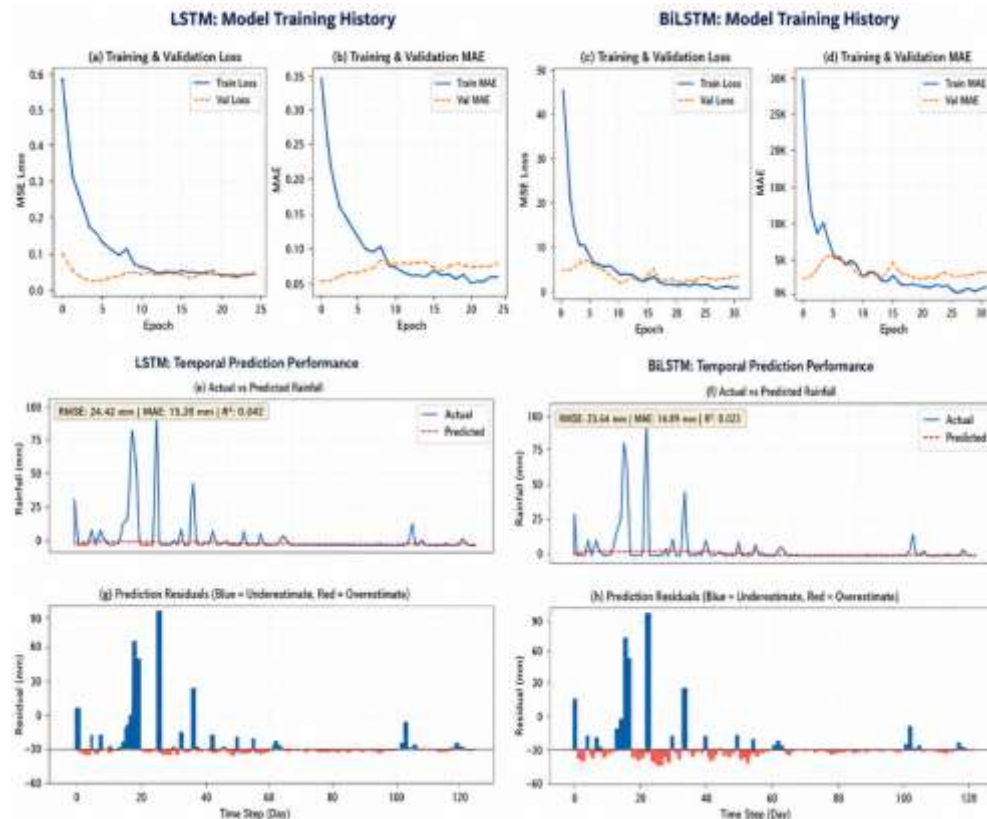


Figure 3. Training History and Prediction Performance of the LSTM and BiLSTM Models on the BMKG Rainfall Dataset

Table 2. Performance Comparison of the LSTM and BiLSTM Models Based on RMSE, MAE, MSE, and R^2 Metrics

Model	RMSE (mm)	MAE (mm)	R^2	MSE (mm^2)
LSTM	24.4181	15.2000	0.0420	596.2456
BiLSTM	23.6398	14.8900	0.0233	558.8401

The results presented in Table 2 indicate that both LSTM and BiLSTM achieved limited predictive performance on the BMKG rainfall dataset. BiLSTM obtained lower RMSE, MAE, and MSE values than LSTM, with RMSE of 23.64 mm, MAE of 14.89 mm, and MSE of 558.84 mm^2 , compared with 24.42 mm, 15.20 mm, and 596.25 mm^2 for LSTM.

The R^2 values were 0.0420 for LSTM and 0.0233 for BiLSTM. Both values are close to zero, indicating that neither model explains a substantial proportion of the observed rainfall variability. Therefore, the lower error values obtained by BiLSTM should not be interpreted as evidence of strong forecasting capability.

The differences between the two models are relatively small. BiLSTM reduced RMSE, MAE, and MSE by approximately 3.19%, 2.04%, and 6.27%, respectively, relative to LSTM, while LSTM produced the slightly higher R^2 value. These differences should be interpreted cautiously because the models were evaluated on a single test split.

Overall, the results show a modest error-based advantage for BiLSTM within the evaluated dataset, but neither model demonstrates strong explanatory capability. Additional analysis using longer historical records, baseline forecasting models, repeated training, and time-series validation would be required before making broader claims about model superiority.

Prediction Analysis

Table 2 indicates that the two architectures generated broadly comparable forecasting errors. The BiLSTM model achieved lower RMSE, MAE, and MSE values, suggesting smaller errors according to these three metrics, while the LSTM model reported a slightly higher R^2 . The gap in R^2 is noteworthy because it shows that the BiLSTM's lower error values did not correspond to a greater share of explained rainfall variability.

*name of corresponding author



This is anCreative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

The R^2 values of 0.0420 and 0.0233 show that both models account for only a small portion of the variability in the observed rainfall series. This finding indicates that rainfall variability in the study area is not sufficiently captured by the current univariate input and the restricted observation period.

Because the experiment relied on a single chronological train-validation-test split and did not include repeated runs or time-series cross-validation, the differences observed here cannot be taken as evidence that the BiLSTM model is generally more capable or more stable than the LSTM model. The correct interpretation is that the BiLSTM model achieved marginally better performance on the error-based metrics for this specific test set.

Accordingly, the prediction analysis supports only a limited comparison within the dataset. The results offer no indication that increasing architectural complexity by itself substantially enhances rainfall forecasting performance, especially given that the available dataset is relatively short.

Error Analysis

The error metrics provide complementary information about the forecasting behavior of the two models. BiLSTM produced an RMSE of 23.64 mm and an MAE of 14.89 mm, while LSTM produced 24.42 mm and 15.20 mm, respectively. The lower RMSE and MAE indicate that BiLSTM had smaller average and root-mean-square errors on the test set.

The difference between RMSE and MAE also indicates that larger residuals contribute substantially to the overall error for both models. However, the current analysis does not contain residual-level or event-level results that would allow the magnitude of errors during specific extreme-rainfall events to be quantified. Therefore, the influence of extreme rainfall should be treated as a limitation requiring additional analysis rather than as a demonstrated model characteristic.

The R^2 results further indicate limited explanatory capability: 0.0420 for LSTM and 0.0233 for BiLSTM. Thus, although BiLSTM achieved lower error-based metrics, LSTM retained the slightly higher R^2 . No single metric provides sufficient evidence to declare one architecture universally superior.

A more informative error evaluation would separate observations according to rainfall intensity, examine residual distributions, and evaluate performance during high-rainfall events. Such analyses were not performed in the current experiment and should be included in future work, especially because rainfall forecasting performance can vary substantially across different rainfall regimes.

Implications for Rice Planting Season Optimization

Beyond comparing prediction accuracy, the rainfall forecasts generated by the LSTM model, which retained the higher R^2 among the two architectures, were further translated into practical implications for rice planting season optimization in Deli Serdang Regency. Following the cropping-calendar approach commonly used by BMKG and the Ministry of Agriculture, monthly rainfall levels were classified into wet, normal, and dry categories, which in turn indicate whether the onset of the rice planting season should be advanced, maintained, or delayed. Months in which predicted rainfall exceeds approximately 200 mm are generally associated with sufficient water availability for land preparation and early-stage rice growth, whereas months with rainfall below 100 mm indicate a higher risk of water deficit that may warrant postponing planting or adjusting irrigation.

Because both models produced R^2 values close to zero and no baseline, repeated-run, or time-series cross-validation was performed in this study, these planting-season indications should be regarded as exploratory rather than definitive. They are best used to complement, not replace, the operational cropping calendar issued by BMKG and the Ministry of Agriculture. Nevertheless, this initial translation of LSTM-based rainfall forecasts into a cropping-calendar classification illustrates a practical pathway toward more adaptive, data-driven rice planting season decision support in Indonesia, provided that future models achieve stronger explanatory power and more rigorous validation before being adopted operationally.

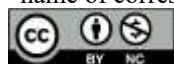
DISCUSSIONS

The main finding is that the difference between LSTM and BiLSTM is modest. BiLSTM achieved lower RMSE, MAE, and MSE, whereas LSTM achieved a slightly higher R^2 . Both models nevertheless produced R^2 values close to zero, indicating limited explanatory capability on the available test data.

Several factors may contribute to the limited performance, including the relatively short observation period, the use of rainfall as the principal time-series input, and the absence of additional meteorological predictors. The 788 observations cover January 2024 to March 2026, which is insufficient to establish strong interannual generalization or climatological behavior.

The comparison should therefore be interpreted strictly as a within-study comparison. Because the experiment did not include persistence, seasonal-naïve, historical-mean, ARIMA/SARIMA, or machine-learning baselines, it cannot establish whether either deep learning architecture improves forecasting accuracy over simpler alternatives. Similarly, because repeated training and time-series cross-validation were not performed, the results do not support claims of model stability or robustness.

*name of corresponding author



This is anCreative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

BiLSTM's lower error-based metrics suggest a modest advantage for the evaluated test set, but the slightly lower R^2 indicates that this advantage is not consistent across all evaluation criteria. The results consequently do not justify a broad claim that BiLSTM is the best architecture for rainfall forecasting in the study area.

The short observation period also limits the assessment of seasonal and extreme-event behavior. A longer historical record would allow the analysis to distinguish recurring seasonal patterns from year-specific variability and would provide a stronger basis for evaluating generalization across different rainfall regimes.

Overall, the evidence supports a cautious conclusion: BiLSTM produced slightly lower error values than LSTM on the available test set, but both models had limited explanatory performance. Future evaluation should combine longer historical data, additional meteorological predictors, baseline models, repeated runs, and time-series cross-validation before practical forecasting benefits are claimed.

These considerations also bear directly on the rice planting season implications discussed above: because the explanatory power of both models remains low, the resulting wet/normal/dry classification should be treated as a preliminary, indicative signal rather than a validated decision tool, and any operational use should await longer historical records, baseline comparisons, and rigorous validation of the underlying rainfall forecasts.

CONCLUSION

This study compared LSTM and BiLSTM for daily rainfall forecasting using 788 BMKG rainfall observations from Deli Serdang Regency collected between January 2024 and March 2026. BiLSTM achieved lower RMSE, MAE, and MSE than LSTM, whereas LSTM achieved a slightly higher R^2 . However, both R^2 values were close to zero, indicating that neither model adequately explains the observed rainfall variability.

The findings should not be interpreted as evidence that either model can, on its own, reliably determine rice planting seasons. While the LSTM-based forecasts were translated into an indicative wet/normal/dry cropping-calendar classification as a preliminary step toward rice planting season optimization, the present study does not include a crop-growth model, a crop-water requirement calculation, or a formal decision-optimization algorithm, and the classification was not validated against actual planting outcomes. In addition, no baseline forecasting models were evaluated, so the study cannot determine whether LSTM or BiLSTM improves upon simpler forecasting approaches.

The study is further limited by the short observation period and incomplete reporting of several experimental details, including sequence length, forecast horizon, missing-value treatment, scaler fitting, final hyperparameters, early stopping, and random seed. Repeated model training and time-series cross-validation were also not performed; therefore, claims regarding robustness, stability, or general superiority are not supported by the current experiment.

Future research should use longer historical records, incorporate additional meteorological predictors, compare the models with persistence, seasonal-naïve, historical-mean, ARIMA/SARIMA, and other relevant machine-learning baselines, and evaluate performance using repeated runs or time-series cross-validation. Residual distributions and performance across rainfall-intensity groups and extreme events should also be examined. Building on the indicative cropping-calendar classification introduced in this study, future work should develop a validated decision layer that combines rainfall forecasts with agronomic information and an explicit, tested decision or optimization rule for rice planting recommendations.

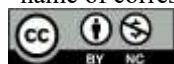
ACKNOWLEDGMENT

The authors would like to express their sincere gratitude to the Indonesian Meteorological, Climatological, and Geophysical Agency (BMKG) for providing the meteorological dataset used in this study. The authors also acknowledge the support of Universitas Senior Medan, particularly the Faculty of Technology and Health Sciences, for facilitating this research. The constructive comments provided by colleagues and anonymous reviewers are greatly appreciated and have substantially improved the quality of this manuscript. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

REFERENCES

- Akbar, A. A., Darmawan, Y., Wibowo, A., & Rahmat, H. K. (2024). Accuracy assessment of monthly rainfall predictions using seasonal ARIMA and long short-term memory (LSTM). *Journal of Computer Science and Engineering*, 5(2), 99–114. <https://doi.org/10.36596/jcse.v5i2>
- Allawi, K. H., & Obayes, H. (2025). *Rainfall forecasting using LSTM*. ResearchGate. <https://www.researchgate.net/publication/388177574>
- Azi, A., & Kusriani. (2025). Rainfall prediction using LSTM in Indonesia. *Journal of Computer Networks, Architecture and High Performance Computing*. <https://jurnal.itscience.org/index.php/CNAPC/article/view/5506>
- Chamatidis, I., et al. (2023). Short-term rainfall forecasting using LSTM. *Environmental Sciences Proceedings*, 26(1), 157. <https://www.mdpi.com/2673-4931/26/1/157>

*name of corresponding author



This is anCreative Commons License This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

- Chang, X. (2024). Rice yield prediction based on deep learning. *Frontiers in Artificial Intelligence and Applications*. <https://doi.org/10.3233/FAIA231333>
- Jabed, M. A., et al. (2024). Crop yield prediction using ML and DL: A review. *Heliyon*. <https://www.researchgate.net/publication/386263273>
- Jumadi, J., Danardono, D., Roziaty, E., Ulinuha, A., Supari, S., Choy, L. K., ... Nawaz, M. (2025). AI-driven ensemble learning for spatio-temporal rainfall prediction in the Bengawan Solo River Watershed, Indonesia. *Sustainability*, 17(20), 9281. <https://doi.org/10.3390/su17209281>
- Kaplun, D., Deka, S., Bora, A., Choudhury, N., Basistha, J., Purkayastha, B., ... Misra, D. D. (2024). An intelligent agriculture management system for rainfall prediction and fruit health monitoring. *Scientific Reports*, 14(1), 512. <https://doi.org/10.1038/s41598-023-49186-y>
- Li, H., et al. (2024). Rainfall prediction using BiLSTM networks. *Water*, 16(2), 245. <https://www.mdpi.com/2073-4441/16/2/245>
- Necesito, I. V., Kim, D.-H., Bae, Y. H., Kim, K., Kim, S., & Kim, H. S. (2023). Deep learning-based univariate prediction of daily rainfall: Application to a flood-prone, data-deficient country. *Atmosphere*, 14(4), 632. <https://doi.org/10.3390/atmos14040632>
- Nemade, B., et al. (2024). Improving rainfall prediction accuracy using an LSTM-driven model enhanced by M-PSO optimization. *Journal of Electrical Systems*, 19(3), 164–180. <https://doi.org/10.52783/jes.664>
- Patro, B. S., et al. (2025). Collaborative deep learning for rainfall forecasting. *Atmosphere*, 16(10), 1197. <https://www.mdpi.com/2073-4433/16/10/1197>
- Peeriga, R. R., Rinku, D. R., Bhaskar, J. U., Nagalingam, R., Aldosari, F. M., Albarakati, H. M., ... Jaffar, A. Y. (2024). Real-time rain prediction in agriculture using AI and IoT: A bi-directional LSTM approach. *Engineering, Technology & Applied Science Research*, 14(4), 15805–15812. <https://doi.org/10.48084/etasr.8011>
- Simanullang, M. J., Rosnelly, R., & Riza, B. S. (2025). Classification of big data stunting in North Sumatra using support vector regression method. *Indonesian Journal of Artificial Intelligence and Data Mining*.
- Tanvir, M. A., & Naiema, N. (2025). Rainfall prediction using machine learning: LSTM. *Journal of Computer Science and Information Technology*, 2(2), 23–33. <https://doi.org/10.61424/jcsit.v2i2.531>
- Wang, F., Cao, Y., Wang, Q., Zhang, T., & Su, D. (2023). Estimating precipitation using LSTM-based raindrop spectrum in Guizhou. *Atmosphere*, 14(6), 1031. <https://doi.org/10.3390/atmos14061031>
- Waqas, M., Humphries, U. W., Hlaing, P. T., & Ahmad, S. (2024). Seasonal WaveNet-LSTM: A deep learning framework for precipitation forecasting with integrated large scale climate drivers. *Water*, 16(22), 3194. <https://doi.org/10.3390/w16223194>