Journal of Artificial Intelligence and Metaheuristic (JAIM)

♦ ISSN: 2833-5597 Vol: 08, No. 01, 2024 ♦ PP. 42–48

Article DOI: https://doi.org/10.54216/JAIM.080105



FULL LENGTH ARTICLE

Weather Prediction: Predicting Rain Using Weather Conditions

Khaled Sh. Gaber[1,*] **Mohamed Abd Elmonem Elsebaey**[1] **Ahmed Al-Sayed Ibrahim**[1]

[1] Department of Communications and Electronics, Delta Higher Institute of Engineering and Technology, Mansoura, 35111, Egypt

Emails: CH1900228@dhiet.edu.eg; CH1900144@dhiet.edu.eg; CH1800067@dhiet.edu.eg

Received: March 22, 2024 Revised: May 20, 2024 Accepted: August 19, 2024 ⋆ Corresponding author

ABSTRACT

Weather forecasting is a major discipline that plays an important role in fields such as agriculture, transport, and emergency management, and it largely depends on accurate forecasts. Concerning this problem, this work aimed to analyze the effectiveness of recurrent neural networks, particularly the Long Short-Term Memory (LSTM), for estimating rainfall depending on precipitation, maximum temperature, minimum temperature, and wind speed. We will therefore use a large database containing recorded weather data obtained over several years to calibrate accurate predictive models designed to distinguish between drizzle, rain, sun, snow, and fog. The main idea of the work is to teach LSTM models that are capable of revealing temporal relations and patterns in sequential data, which makes them suitable for time-series forecasting such as weather prediction. The data is preprocessed effectively to clean it and make it suitable for analysis. To accurately compare the performance of one model against the others, we have divided the data into training, validation, and testing sets. The performance of the proposed LSTM model is evaluated with the Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and the coefficient of determination (R^2) to measure forecasting accuracy. The findings show improved predictive performance: the best-performing LSTM model has an MSE of 8.74, RMSE of 2.96, MAE of 2.35, and R^2 of 0.83. These metrics indicate a relationship between the predicted and actual weather conditions, demonstrating the model's efficiency. The evaluation also shows how hyperparameter optimization, feature selection, and normalization affect model performance and how precise management of weather parameters can result in better forecasts. This research has practical as well as theoretical value because it can improve the dependability of weather forecasts. More precise weather data can benefit crop growing, road networks, transportation systems, and emergency and rescue operations preparing for adverse conditions and disasters. Consequently, this study contributes to research on weather patterns and unexpected rainfall by demonstrating the potential of LSTM networks to analyze key meteorological characteristics for rainfall prediction. Possible future directions include expanding the feature set, using continuous rather than weekly data, and considering a mixed-input approach to increase prediction accuracy. This strategy seeks to improve meteorological forecasting and timely resource allocation and management under climate volatility.

Keywords: Weather Prediction ▪ Rainfall ▪ LSTM ▪ Machine learning ▪ Meteorological variables ▪ Time-series forecasting

1. INTRODUCTION

Weather forecasting is important in many fields of modern life, including agriculture, transport, and natural-disaster management. Weather forecasts help users make decisions about their operations, enhancing efficiency and reducing costs. Forecasts are now necessary for activity planning because of unstable weather conditions resulting from global warming. Conventional prediction models may have difficulty handling

© 2024 The Author(s). Published by ASPG. This article is licensed under CC BY 4.0.

the large amounts of nonlinear information inherent in meteorological data [1, 2].

Further developments in machine learning, especially deep learning, seem promising for increasing the reliability of weather forecasts (Hoi and Lam, 2015; LSTM, 2016). LSTMs are designed to handle temporal patterns in sequential data, making them suitable for time-series prediction. Because these models can train on past weather conditions, they can capture finer details and dependencies that other methodologies may not be able to [3].

In this study, we predict rainfall using weather data such as rainfall, maximum temperature, minimum temperature, and wind speed. These features are important for defining weather outcomes, and their inclusion in prediction models provides a fuller picture of weather conditions. The main goal of the LSTM-based model is to classify several types of weather, including drizzle, rain, sun, snow, and fog.

The concern of this research is not limited to academic purposes: reliable weather prediction matters to individuals interested in climatic patterns. Planning agricultural activities based on meteorological information can increase yields, while using such information to assess transportation safety can improve efficiency. Emergency services can also make arrangements in anticipation of impending disasters to save lives and property.

2. LITERATURE REVIEW

Crop diseases are well known to cause substantial losses to yields in agricultural production. This research uses Convolutional Neural Networks (CNNs) to diagnose crop diseases from images, with the model operating on leaf images to determine whether they are infected. The proposed system is more accurate than conventional image-processing techniques for classifying and detecting diseases at an early stage. The developed CNN model has the potential to help farmers analyze conditions affecting crops using training data of diseased and healthy leaves [1].

The Internet of Things (IoT) is a revolutionary technology that can help in farming and other sectors by allowing monitoring devices and technologies to feed real-time data. The research proposes an IoT architecture for precision farming to manage crop growth and resources through sensors for environmental and soil conditions. Machine-learning models used in the system help farmers in their decision-making process through trend identification to predict the best times to water crops, early identification of problematic crop-growth patterns, and weather forecasting [2].

Drones are increasingly used in contemporary agriculture for precision surveillance and data acquisition. This research examines drones that incorporate multispectral cameras to assess crop health and plant vigor and to screen for diseases early. When used to fly over fields and capture images and video, drones can help determine soil variability, nutrients required for plant growth, and stress factors affecting field yields [3].

Higher demand for global food production leads to advancement in smart agricultural systems. This research presents a smart-greenhouse management system using IoT technology with a machine-learning algorithm. It controls external environmental parameters for plants, including temperature, humidity, and carbon dioxide concentration, and regulates them at the necessary levels with minimal time delays. A machine-learning algorithm is trained using previous climate data to determine the best climate setting needed for the greenhouse and to make environmental control independent, achieving optimum crop yields [4].

Machine learning can predict crop yields, and this data is useful for food security and production strategy. This research measures the performance of machine-learning algorithms, including random forests, support-vector machines, and artificial neural networks, for estimating the yields of crops influenced by soil type, weather variables, and fertilizer application. The study establishes the superiority of the random-forest model for crop-yield prediction systems and its ability to enhance prediction accuracy [5].

Climate change threatens agricultural productivity across the world. This research explores the impact of temperature and precipitation changes on plant germination, growth, and productivity, with the aim of developing models for effective crop management under changing climatic conditions. Using climate information about weather conditions together with crop models gives guidance on how climate conditions might affect future crop production and shows possible ways to address those effects by changing planting dates and crop-management systems [6].

Precision-agriculture technologies maintain the efficiency of resource use and enhance production output. The authors use machine-learning algorithms to develop a model for determining the right fertilizer application by considering soil and crop features. The study presents a decision model focused on choosing the correct fertilizer type and amount to minimize waste and harm to the environment and promote plant development. Field trials show that the system helps increase nutrient management on farms [7].

Weeds are a threat to modern agriculture, so detection and control should be performed automatically. This paper provides a deep-learning framework for weed detection in crop fields by feeding images from field cameras through convolutional neural networks. The model distinguishes crops from weeds, and herbicide is applied only to the weeds, reducing chemical use and harm to crops. The performance demonstrated in the study suggests the ability to detect weeds irrespective of lighting and growth conditions [8].

Water scarcity is an important factor in the agricultural subsector, especially in arid regions. This work suggests an artificial-intelligence approach to estimate needed irrigation given soil-moisture status, meteorological conditions, and crop water requirements. By combining sensor data with models, the system can determine when irrigation is required while saving water and providing crops with the necessary amounts. The study shows that water use can be optimized in irrigation systems through artificial intelligence, primarily machine learning [9].

Nutrient deficiencies in plants cause drastic decreases in yield. This research proposes developing a machine-learning system for identifying nutrient deficiencies in plants from leaf images. The approach uses color changes in leaf images to detect deficiencies of nitrogen, phosphorus, potassium, and other

nutrients. A farmer may fail to identify nutrient deficiencies until they have severely affected productivity; by detecting nutrient issues early, farmers are better positioned to tackle them before they worsen [10].

Sentiment analysis is the capability to predict the sentiment polarity of a given piece of text as positive, negative, or neutral. It is practiced in natural-language processing, computational linguistics, and data mining. Arabic presents complexity because of its inherent ambiguities and dialects, as well as the paucity of resources. Research reports high performance with BiLSTM as the model of choice, which effectively captures contextual information. This paper focuses on applying BiLSTM to enhance Arabic sentiment analysis because it can capture contextual information from Arabic feature sequences and improve performance over previously introduced machine-learning techniques [11].

Recurrent Neural Networks are typically used to handle sequential data such as language, text, audio, and video. However, ordinary RNNs have long-term dependency problems. The cell structure with gate functions in the Long Short-Term Memory (LSTM) network addresses this drawback and is therefore an object of deep-learning study. This paper briefly introduces LSTM cells and their types, categorizes LSTM networks, discusses their application areas, and provides prospects for subsequent research on LSTM networks [12].

Reference evapotranspiration (ET_0) must be predicted accurately to determine the amount of water needed for crop irrigation. This analysis employs LSTM and BiLSTM models for forecasting daily ET_0 rates and multi-step prediction. LSTM showed the highest accuracy when predicting daily total amounts, while BiLSTM was more efficient for forecasting multi-step totals. The outcome established that BiLSTM models incorporate flexibility across diverse datasets pertinent to ET_0 prediction [13].

Floods and droughts are complex spatiotemporal events affected by weather and physical geography. This research proposes a forecasting system based on an LSTM feedforward neural network for hydrological extreme events in the South Pacific, derived from satellite rainfall and sea-surface-temperature anomalies. Extreme-event detection is enhanced by including eigenvector values from SST data rather than using only the basic univariate and multivariate framework. Furthermore, the model's multivariate approach provides research hypotheses for future application in long-range weather prediction [14].

Artificial intelligence and machine learning are increasingly applied to digital marketing. One study examined an ensemble approach for food-delivery businesses, combining decision trees, naive Bayes, and nearest-neighbor methods to analyze customer preferences and predict behavior for personalized recommendations [15].

3. PROPOSED METHODOLOGY

This section provides an overview of the systematic approach used for weather prediction with selected machine-learning methods. Data collection and preprocessing are the first steps, intended to provide clean, high-quality data. The predictors are past weather records, with features including precipitation, maximum and minimum temperature, and wind speed.

To prepare the data for modeling, missing values are handled and feature scales are normalized. Feature selection then identifies the most influential predictors of weather conditions. Long Short-Term Memory networks are used because they are suitable for sequential data and temporal characteristics.

The methodology also comprises model training, evaluation, and validation using performance indicators such as Mean Squared Error (MSE) and Root Mean Squared Error (RMSE). The metrics are presented in a formatted manner to make the developed model's performance clear. The optimized routine proposed in this work is intended to improve the precision of meteorological findings and contribute to the development of weather-prediction methods.

3.1 Dataset Description

The data used in this research include historical weather records obtained from meteorological databases spanning different years for seasonality analysis. They include key features for predicting weather conditions:

- **Precipitation:** Expressed in millimeters (mm), this feature is used to predict rain, drizzle, and snow patterns that reach the ground.
- **Maximum temperature (temp_max):** The highest temperature of the day, expressed in degrees Celsius (°C), affecting evaporation rates and general weather.
- **Minimum temperature (temp_min):** Expressed in degrees Celsius (°C), this is the lowest temperature observed during the day, usually at night.
- **Wind speed:** Given in km/h or m/s, wind speed affects the rate and regularity of precipitation and weather stability.
- **Weather-condition labels:** Categorical labels including drizzle, rain, sun, snow, and fog; these are the target variable to be predicted.

The available dataset contains thousands of daily readings, and each record includes the date and relevant weather characteristics. Preprocessing steps were conducted to improve data quality, including cleaning missing values and outliers. This comprehensive database is intended to support the creation of a machine-learning algorithm for weather prediction and provide information relevant to meteorology.

3.2 Descriptive Analysis

This section conducts exploratory analysis of the weather dataset to establish possible relationships and distributions among variables. Knowledge of these relations is important for building prediction models. Characteristics of interest include maximum and minimum temperatures, rainfall, and wind speed over a given period.

Figure 1 is a line graph labelled "Temperature Max & Min" and depicts maximum and minimum temperatures from January 2012 to January 2016. The horizontal axis shows dates from January 2012 to January 2016, and the vertical axis shows temperature in central Sweden, generally between -10 and 40 degrees Celsius. The blue line represents maximum temperature and the orange line represents minimum temperature.

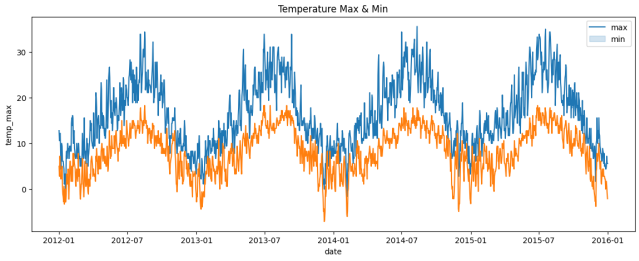


Figure 1. Temperature Max & Min.

The line graph shows that maximum temperatures are highest in the middle of the year, indicating summer, and lowest near the end or beginning of the year, indicating winter. In particular, TMAX consistently exceeds TMIN throughout the observed period, as expected with seasonal climate changes.

The second visualization, Figure 2, is a pair plot that explores relationships among four key weather variables: rainfall, maximum temperature, minimum temperature, and wind speed. The scatter plots in the matrix show relationships between pairs of variables, while the diagonal illustrates the distribution of each variable. The coordinates are color-coded by weather type: blue for drizzle, red for rain, green for sun, purple for snow, and orange for fog.

This visualization is useful for showing trends and relationships among aspects of weather that depend on several factors. For example, high precipitation is associated with low maximum and minimum temperatures during rainy and snowy weather. This can help build the weather-prediction model and identify potential predictors with the pair plot. Descriptive statistics provide an understanding of the dataset for the subsequent modeling stage. These visualizations show key trends, and understanding those trends is an essential first step in modeling relationships between weather conditions.

3.3 Data Preprocessing

Cleaning and transforming the weather dataset are important for structuring the data for analysis. Several processing methods were used to improve the quality of the dataset, including its precipitation, maximum-temperature, minimum-temperature, and wind-speed features [4].

1. **Handling missing values:** Initial preprocessing filled missing values in the dataset. Gaps in numerical features, including temperature and wind speed, were addressed by replacing missing values with nearby values to preserve the dataset's structure. Rows with more missing values than the established limit would be excluded.

2. **Normalization and scaling:** Normalization was applied to feature values so that the model could train effectively. Feature ranges varied from millimeters of precipitation to degrees Celsius, so Min–Max scaling was used to scale features to the $[0,1]$ range. This prevents features with large ranges from dominating the learning process.

3. **Feature engineering:** Feature engineering was important for improving predictive accuracy. It gave additional meaning to existing features related to climate. For instance, average temperature was calculated as the arithmetic mean of the highest and lowest temperature readings, providing a broader view of temperature fluctuations for climate estimation.

4. **Data splitting:** After preprocessing, the database was divided into training, validation, and test data. The training set includes 80% of records and is used for model training; the validation set contains 10% and is used to test model performance and tune hyperparameters; and the final 10% is used as the test set to assess predictive accuracy independently of training.

Table 1. Criteria for evaluating regression results.

Metric	Formula
Mean Squared Error (MSE)	$\frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2$
Root Mean Squared Error (RMSE)	$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}$
Mean Absolute Error (MAE)	$\frac{1}{n}\sum_{i=1}^{n}\lvert y_i-\hat{y}_i\rvert$
Coefficient of determination (R^2)	$R^2 = 1-\frac{\text{SSE}}{\text{SST}}$

3.4 Artificial Neural Networks (ANN)–LSTMs

LSTM is a recurrent neural-network architecture proposed by S. Hochreiter and J. Schmidhuber in 1997 [5]. It uses a gated mechanism (input, output, and forget gates) to regulate the flow, storage, and dependencies of information during sequence training. Long Short-Term Memory eliminates long-term dependency, vanishing-gradient, and exploding-gradient issues in traditional recurrent neural networks. LSTM networks can handle long sequences using gates that regulate interactions with inputs [6]. An LSTM cell unit consists of input, output, and forget activation vectors [6].

LSTM is better suited to noisy and nonlinear data, unlike methods such as ARIMA, which handle linear time-series prediction. Prior work reports that LSTM is superior for large-scale data, while ARIMA may be superior for smaller-scale data because of inadequate sample sizes [7]. Other models, such as LSTM-RNN, have been deployed with higher accuracy than other models [8]. Here, LSTM is used to capture temporal weather patterns and make predictions by learning long-term dependencies with memory units [9, 10, 11, 12].

3.5 Performance Metrics

The performance of the models developed in this study was measured using statistical evaluation criteria [13]. The same model was assessed on training data and an independent test set [14]. This study computed RMSE, MSE, and R^2 at three time intervals, with averages obtained from 30 repeated trials. RMSE for each grid cell across the training and testing phases was also calculated.



Figure 2. Pair plot of weather data.

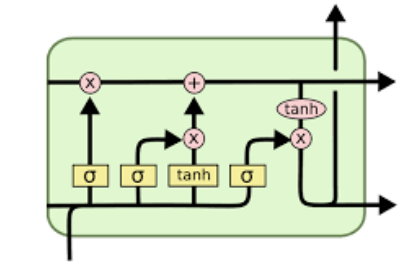


Figure 3. LSTM architecture.

$$\text{SSE} = \sum_{i=1}^{N} (y_i^{\text{actual}} - y_i^{\text{predicted}})^2,$$

$$\text{SST} = \sum_{i=1}^{N} (y_i^{\text{actual}} - \bar{y}^{\text{actual}})^2.$$

- **Mean Squared Error (MSE):** This is the mean of the squared differences between forecast and actual values. It is used to test the model's prediction error; lower MSE values indicate a better model [1].
- **Root Mean Squared Error (RMSE):** RMSE is the square root of MSE and directly measures the average magnitude of prediction errors. A low RMSE is better [1].
- **Mean Absolute Error (MAE):** MAE estimates the average size of errors and represents their range [1].
- R^2 **(coefficient of determination):** R^2 measures the variance in the dependent variable accounted for by the independent variable. Higher R^2 values correspond to better model fitting [1].

4. RESULTS

The model's performance was measured using Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and the coefficient of determination (R^2). These metrics provide insight into how accurately the model predicts results on test data. The results are an MSE of 8.74, RMSE of 2.96, MAE of 2.35, and R^2 of 0.83. These values indicate that the model makes accurate predictions, establishes a relationship between predicted and actual values, and has relatively low error rates.

Table 2. The model results.

MSE	RMSE	MAE	R^2
8.741861	2.956664	2.353954	0.827729

The actual and predicted values for the validation and test datasets are illustrated in Figure 4. The figure shows how closely the model captures actual values and trends, with little discrepancy between actual and predicted values. The correlation is high, indicating generalization to unseen data. Small variations in some areas could be explained by factors outside the features considered. These deviations remain within reasonable limits and do not seriously affect the quality of the resulting predictions.

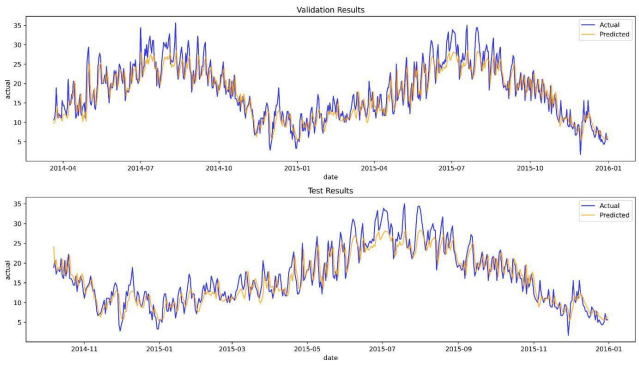


Figure 4. Comparison of predicted and actual values.

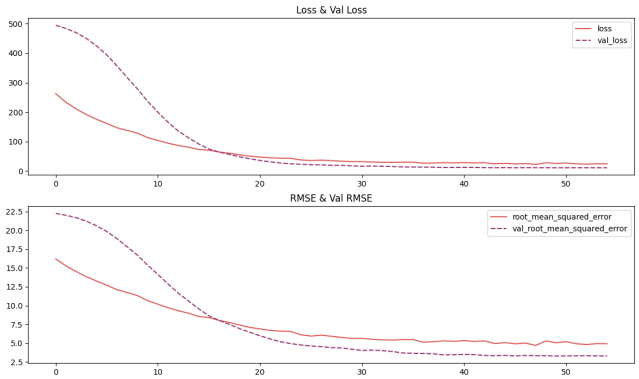


Figure 5. Training and validation loss curves.

The training process is shown in Figure 5 in terms of loss and RMSE throughout the epochs. The reduction in training and validation losses indicates that the model is learning from the data. Validation loss remains slightly higher than training loss, as usual, but the difference is small, indicating that the model did not overfit. The progressive decline and continuous fluctuation of the loss suggest convergence and the capacity to perform well with other datasets.

The evaluation metrics and visualizations show that the model is efficient in making precise predictions with comparatively small error rates. The high R^2 value (0.83) reflects the proportion of target-data variance captured by the model. The values indicate that the model has captured significant data characteristics and applies them to unknown data points with small discrepancies between actual and estimated values. The difference between predicted and actual values shown in Figure 4 is negligible; adding features or adjusting hyperparameters may reduce it further. The training and validation loss trajectories in Figure 5 show that the model learned properly and did not overfit. These results suggest that the model is reliable for downstream analyses and can capture target effects, assuming finer-grained feature extraction or hyperparameter optimization. Figure 4 compares predicted and actual values, while Figure 5 illustrates the loss and RMSE curves.

5. CONCLUSION

The purpose of this paper was to research how climate variables and climate models can be included in a model to improve recommendations for crop yield using different agricultural inputs. Evaluation of the proposed model suggested good performance based on Mean Squared Error (MSE) of 8.74, Root Mean Squared Error (RMSE) of 2.96, and a coefficient of determination (R^2) of 0.83. These results show that the model captures climate-dependent variables, soil conditions, and crop yields and generates realistic predictions that can assist farmers and stakeholders in agriculture. The findings are illustrated by the predicted and actual values (Figure 4) and the training and validation loss (Figure 5). The small discrepancy between predicted and actual data demonstrates the model's ability to generalize to unseen data, while the loss curves indicate learning without overfitting. This balance between training and validation loss is important for confidence in applying the model in agricultural practice.

Nevertheless, the work could be extended. Future research may address more detailed climate data and additional factors, such as weather change and storms, to improve predictive ability. More sophisticated machine-learning methods and extensive hyperparameter optimization may also improve accuracy and speed. This work advances knowledge in climate data, forecasting, and their application within crop-recommendation systems. Through the use of big data, actors in the farming value chain can better select crops, improve adaptability to climate change, and increase food productivity. These insights highlight the need for advanced technologies to improve farming mechanisms likely to be affected by climate change.

REFERENCES

[1] J. A. Segovia, J. F. Toaquiza, J. R. Llanos, and D. R. Rivas, "Meteorological variables forecasting system using machine learning and open-source software," *Electronics*, vol. 12, no. 4, p. 1007, 2023.

[2] P. Aghelpour, B. Mohammadi, and S. M. Biazar, "Long-term monthly average temperature forecasting in some climate types of iran, using the models SARIMA, SVR, and SVR-FA," *Theoretical and Applied Climatology*, vol. 138, no. 3, pp. 1471–1480, 2019.

[3] R. Llugsi, S. E. Yacoubi, A. Fontaine, and P. Lupera, "Comparison between Adam, AdaMax and AdamW optimizers to implement a weather forecast based on neural networks for the andean city of quito," in *2021 IEEE Fifth Ecuador Technical Chapters Meeting (ETCM)*, 2021, pp. 1–6.

[4] E. Alshdaifat, D. Alshdaifat, A. Alsarhan, F. Hussein, and S. M. F. S. El-Salhi, "The effect of preprocessing techniques, applied to numeric features, on classification algorithms' performance," *Data*, vol. 6, no. 2, p. 2, 2021.

[5] M. A. R. Suleman and S. Shridevi, "Short-term weather forecasting using spatial feature attention based LSTM model," *IEEE Access*, vol. 10, pp. 82 456–82 468, 2022.

[6] D. Kreuzer, M. Munz, and S. Schlüter, "Short-term temperature forecasts using a convolutional neural network—an application to different weather stations in germany," *Machine Learning with Applications*, vol. 2, p. 100007, 2020.

[7] A. Parasyris, G. Alexandrakis, G. V. Kozyrakis, K. Spanoudaki, and N. A. Kampanis, "Predicting meteorological variables on local level with SARIMA, LSTM

and hybrid techniques," *Atmosphere*, vol. 13, no. 6, p. 878, 2022.

[8] R. Zafar, B. H. Vu, M. Husein, and I.-Y. Chung, "Day-ahead solar irradiance forecasting using hybrid recurrent neural network with weather classification for power system scheduling," *Applied Sciences*, vol. 11, no. 15, p. 6738, 2021.

[9] P. Hewage, M. Trovati, E. Pereira, and A. Behera, "Deep learning-based effective fine-grained weather forecasting model," *Pattern Analysis and Applications*, vol. 24, no. 1, pp. 343–366, 2021.

[10] S. Khullar and N. Singh, "Water quality assessment of a river using deep learning Bi-LSTM methodology: Forecasting and validation," *Environmental Science and Pollution Research*, vol. 29, no. 9, pp. 12 875–12 889, 2022.

[11] H. Elfaik and E. H. Nfaoui, "Deep bidirectional LSTM network learning-based sentiment analysis for arabic text," *Journal of Intelligent Systems*, vol. 30, no. 1, pp. 395–412, 2021.

[12] Y. Yu, X. Si, C. Hu, and J. Zhang, "A review of recurrent neural networks: LSTM cells and network architectures," *Neural Computation*, vol. 31, no. 7, pp. 1235–1270, 2019.

[13] D. K. Roy, T. K. Sarkar, S. S. A. Kamar, T. Goswami, M. A. Muktadir, H. M. Al-Ghobari, A. Alataway, A. Z. Dewidar, A. A. El-Shafei, and M. A. Mattar, "Daily prediction and multi-step forward forecasting of reference evapotranspiration using LSTM and Bi-LSTM models," *Agronomy*, vol. 12, no. 3, p. 594, 2022.

[14] A. Anshuka, R. Chandra, A. J. V. Buzacott, D. Sanderson, and F. F. van Ogtrop, "Spatiotemporal hydrological extreme forecasting framework using LSTM deep learning model," *Stochastic Environmental Research and Risk Assessment*, vol. 36, no. 10, pp. 3467–3485, 2022.

[15] C. Yaiprasert and A. N. Hidayanto, "Ai-driven ensemble three machine learning to enhance digital marketing strategies in the food delivery business," *Intelligent Systems with Applications*, vol. 18, p. 200235, 2023.