Forecasting Daily Relative Temperature in Hmawbi, Myanmar Using Convolutional Neural Network and Deep Learning Models

Nge*

Department of Computer Engineering and Information Technology, Technological University (Thanlyin), Yangon, Myanmar

***Corresponding Author**

DOI: https://doi.org/10.47772/IJRISS.2026.100600311

Received: 27 May 2026; Accepted: 01 June 2026; Published: 23 June 2026

ABSTRACT

Relative temperature (maximum and minimum) is a significant factor influencing climate change and global warming, and it has become an important research focus in recent decades. This study aims to apply a deep learning approach using a Convolutional Neural Network (CNN) model to forecast daily relative temperature in Hmawbi Township, Myanmar. The model is developed using average daily relative temperature data collected from 2013 to 2022. Based on the forecasting results, a slightly decreasing trend in relative temperature is observed. The CNN model is employed to learn the underlying patterns in historical temperature data collected from 2013 to 2017 and to generate forecasts for the period 2018–2022. In this paper, two prediction models are developed by using Convolutional Neural Network (CNN): one with the original data (without noise removal), one with data cleaned using the Simple Moving Average method (SMA). The performance and accuracy of the CNN model are evaluated using root mean squared error (RMSE) and mean square error (MSE). The use of CNN allows for effective feature extraction and improved handling of nonlinear relationships in the dataset. CNN achieves RMSE of 5.398 and MSE of 29.020, SMA achieves RMSE of 3.767 and MSE of 14.841. The results indicate that the CNN with SMA deep learning model achieves lower error values, demonstrating higher prediction accuracy and better performance compared to traditional statistical methods. The CNN-based deep learning model is found to be a reliable and effective approach for forecasting relative temperature in Hmawbi Township.

Keywords: Forecasting, Deep Learning, CNN, SMA, Seasonality, Maximum and Minimum Temperature

INTRODUCTION

This study focuses on relative temperature in Hmawbi Township, Myanmar, by utilizing daily maximum temperature and minimum temperature data. By incorporating both maximum and minimum temperature, the model can better represent diurnal temperature changes and improve prediction performance. The main objective of this study is to evaluate the effectiveness of a CNN-based deep learning model in forecasting relative temperature using daily maximum and minimum temperature data. The findings are expected to contribute to more accurate climate prediction and support decision-making in climate-sensitive sectors.

Since the early nineteenth century, people have attempted to forecast weather conditions, initially relying on simple observations and experience. Weather forecasting is the process of predicting atmospheric conditions for a specific location and time using scientific and technological methods. In its early stages, forecasting was performed manually, based on factors such as barometric pressure, cloud patterns, and visible weather changes. Over time, advancements in technology have transformed this process into a data-driven discipline that uses computer-based models and large-scale atmospheric data. Today, modern weather forecasting systems integrate various meteorological variables and apply computational techniques to improve prediction accuracy. Accurate weather forecasts are essential for many sectors, including agriculture, transportation, energy production, disaster management, and environmental monitoring. Reliable predictions support better decision-making, efficient resource management, and enhanced public safety.

Traditional forecasting methods often rely on physical and statistical models, which can struggle to capture complex nonlinear relationships in weather data. Machine learning approaches address this limitation by analyzing historical data to identify patterns and relationships for predicting future conditions. In particular, deep learning techniques have shown significant potential due to their ability to automatically extract meaningful features from large datasets. The CNN model is designed to learn complex temporal patterns from historical temperature data and generate accurate forecasts. Compared to traditional models and other machine learning techniques, CNN offers improved capability in handling nonlinear relationships and large datasets.

The objective of this study is to evaluate the effectiveness of a CNN-based deep learning approach in forecasting relative temperature. By leveraging daily maximum and minimum temperature data, this research aims to enhance prediction accuracy and contribute to more reliable weather forecasting systems for practical applications. This study is to enhance the accuracy and reliability of weather forecasting systems through the application of advanced meteorological data analysis techniques, with a particular focus on deep learning models such as Convolutional Neural Network (CNN). Weather prediction remains a complex task due to the dynamic and nonlinear nature of atmospheric processes.

Through this comprehensive analysis, the study aims to advance weather forecasting systems, improving their precision and practical relevance for applications such as agricultural planning, climate change analysis, and daily decision-making. The remainder of this paper is organized as follows: Section 2 reviews related work in weather forecasting and prediction models. Section 3 describes the background theory with emphasis on CNN architectures. Section 4 discusses the performance of the proposed models. Section 5 presents the experimental results of the system. Finally, Section 6 concludes the paper by summarizing key findings.

Related Works

In this paper [1], the researchers gathered historical weather data from various sources, such as temperature, humidity, wind speed, precipitation, and atmospheric pressure, specific to the Shenzhen region. They then developed a deep learning model, which is a type of artificial neural network, to analyze and learn from this vast dataset. The deep learning model utilizes a combination of convolutional neural networks (CNNs) and recurrent neural networks (RNNs) to capture both spatial and temporal patterns in the weather data. CNNs are effective in recognizing spatial patterns, such as temperature variations across different geographical locations, while RNNs excel in capturing temporal patterns, such as daily or seasonal weather changes. Once the deep learning model is trained, it can make accurate predictions of various weather parameters for the next day in Shenzhen City. These predictions include temperature, humidity, wind speed, precipitation, and other relevant factors. The model considered both the historical data and the current weather conditions to generate forecasts. The study evaluated the performance of the deep learning model by comparing its predictions with actual observed weather data. It measured various statistical metrics, such as mean absolute error (MAE) and root mean square error (RMSE), to assess the accuracy and reliability of the forecasts.

The authors implemented, the Deep Learning-Based Weather Prediction using Long Short-Term Memory (LSTM) which is an artificial Recurrent Neural Network (RNN) architecture [2]. Like numerical weather prediction, deep learning-based weather prediction is quite sensitive to the commencement value. The precision of forecast results is strongly influenced by the quality of the input datasets. Sensors and autonomous observing platforms, spanning ocean-based, ground-based, air-based, and space-based platforms, have accumulated Pressure Barometric (PB) of meteorological observation data.

This paper predicted weather attributes such as minimum temperature, maximum temperature, humidity and wind speed using different deep learning methods, convolutional neural networks (CNN), long short-term memory (LSTM) and ensemble of CNN and LSTM (CNN-LSTM) [3]. Dataset including daily weather conditions from past two decades of Yangon Region, Myanmar was used as case study for this research and one-month weather conditions will be predicted. Root Mean Square Error (RMSE) is used for comparison of the performance of these methods. The experiment described that ensemble CNN-LSTM model has better performance than other individual methods.

The authors described data mining techniques for agricultural meteorological features collected from Bengaluru district's meteorological center [4]. The feature extraction with hierarchical clustering and K-means which play an important role in the decision making for the experimental result showed that the hierarchical approach supports better than K-means and K-means clustering gives effectiveness for prediction of weather.

The authors proposed a system the chances of rainfall forecasting with big data predictive analytics in Hadoop framework [8]. The capturing of relationships among many data factors was performed by this model to place the score or weight pattern for rainfall prediction using historical data. This system provided the effectiveness in huge amount of data processing with big data analytics techniques. In this system, the analytics was performed by classification of type of weather with Naive Bayes and humidity weather attribute. The design of system and the plot of mean, maximum and precipitation parameter of humidity are applied to obtain the weather prediction enhancement.

The author proposed the geostatistical interpolation technique, Kriging, for short-term prediction of weather with various observations of weather with surveillance data [9]. The accurate capturing in the spatial-temporal distribution of wind and temperature could be performed by this approach that supports to achieve high-quality local and short-term prediction of weather with uncertainty measurement. The UK-ST framework had been used to perform the prediction of spatial-temporal variograms based on temperature and wind speeds with the trend models. The cross-validation with many methods were evaluated in order to prove that the accurate capturing in the spatial-temporal distribution of these weather variables can be performed by the generated wind and temperature models.

Background Theory

Deep Learning

Deep learning is a kind of artificial intelligence and machine learning that imitates the path humans take to achieve specific kinds of knowledge [8]. This learning is a potential thing in data science that contains modeling of predictive and statistical methods. This is more efficient for data scientists who are tasked with the analysis, translation, and collection of a huge volume of data; this learning makes this procedure easier and faster. Operational models are permitted by incorporating many operation layers for learning descriptions in the data at many abstraction levels. This learning model contains many fully connected hidden layers, [5]so such models are called "deep learning models" in comparison with models with only two hidden layers, which are called "shallow learning models." The classification of deep neural networks is done according to the flow of information. When the flow of information from the input layer to the output layer is performed without feedback, this network is known as a feedforward deep learning neural network.

Learning descriptions from raw datasets is one of the primary activities of deep learning neural networks. The neural network model has the ability to automatically discover the descriptions in data needed to detect features and classify them; this is called feature learning or description. A deep learning neural network may be specified with a class of machine learning methods that apply many layers of operational elements for continuous learning and extract features from a raw dataset [6]. As the movement from the lower end to the higher end of the layers is done, the extracted features start resulting in more and more pronounced results in the learning model for inferring desired results for the given classification or forecasting job.

Convolutional Neural Network

A convolutional neural network (CNN) is a type of deep, feedforward neural networks. It applies the multilayer perceptron and it has been developed for the reduction of processing needs [9]. It can be applied for the collection of spatial data or sequential data and it is widely utilized in speech recognition, time series analysis and image processing, etc. The architecture is shown in Figure 1.

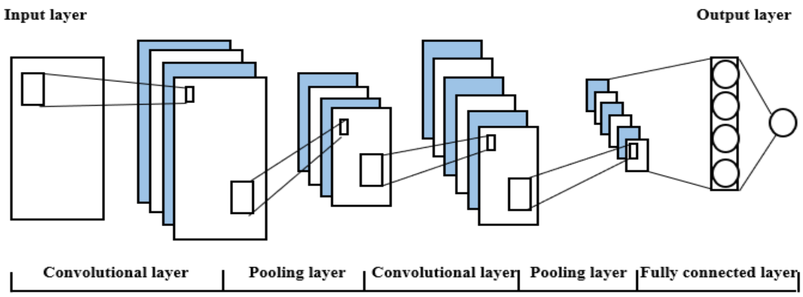

Fig. 1. Convolutional Neural Network

It consists of an input layer, an output layer and a hidden layer which contains many pooling layers, convolutional layers, and fully connected layers.

In Input layer, text documents or images are kept by the representation of vector.

Convolutional layer, the output is decided with the calculation of dot products between set of weights at input layer. The aim of this layer is the observation of features. The principal construction block of the convolutional neural network is the convolutional layer that performs the creation of feature activation maps from the input layer applying many related objects called filters, by passing the certain filter across the height and width of the input layer, therefore, for the calculation of the dot product among the input layer and entries in the filter. The convolution computation occurs in the creation of many two-dimensional activation maps, that are later fed into continuous pooling layers. The times of movement of the filter for each stage is decided by the stride's value, the default is one. The activation function Rectified Linear Units (ReLU), is applied by the convolutional layer for the conversion of all the negative values into value zero. Every convolutional layer is specified by various parameters containing kernel size, zero padding, stride, input size, and the map stack. Then, the calculation of input signal is performed with an activation function, Rectified Linear Units (ReLUs). This activation function is used on the input data; therefore, the dimensions of the input and the output are same.

In Pooling Layer, this layer acts as a mediator between many convolutional layers. It performs the reduction of the spatial dimensions of the input data. This is similar to the previous convolutional layer as this layer sweeps the filtering

among all input data however this filtering does not possess any weights. Two kinds of pooling operations are maximum pooling, and average pooling. Maximum pooling takes the selection of the pixel by the maximum value for sending to the output array when the movement of filter among the input. Average pooling performs the computation of the average value in the receptive field by sending to the output array when the movement of filter among the input. This layer is applied for the down sampling, that reduces the disperse size of created feature maps with convolution layer with the reducing of their dimension according to a selected criterion. Pooling directs to extract the most significant feature from the feature maps and optimize the overall computation required for data processing. This layer provides many benefits to convolutional neural network whereas there is the information loss at this layer. The reduction of noise features, the efficiency improvement, and the overfitting prevention are provided by this layer.

Fully-Connected Layer*:* This layer is same with the output layer of multilayer perceptron (MLP). The aggregation of information from the final feature maps and the generation of final classification are performed. The fully connection of all neurons with all neurons in the previous layer is taken. This is the last element of convolutional neural network and is a fully connected layer of the preceding artificial neural networks. The flattened column vector then becomes the input for the fully connected layer for transforming feature vectors, that is performed by training the network utilizing backpropagation on many epochs. The last element of the fully connected layer applies an activation function like a sigmoid or softmax activation function for the class label forecasting creation, that is also the final result of convolutional neural network [7]. The reduction of data dimension at pooling layer to a single dimension and the connection with every neuron are taken. The classification is performed with activation function, SoftMax.

$$\sigma(\vec{z})i = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}} \qquad (1)$$

where, $\vec{z}$ is the input vector, z_i is the elements of the input vector, e^{z_i} is the standard exponential function, and K is the number of classes in the multi-class classifier. The SoftMax function does the testing of result by using training data. Finally, this provides the result which the input data to which the relating class.

Simple Moving Average

Simple Moving Average (SMA) is a fundamental and widely used time series in weather forecasting systems. It serves as a valuable tool for smoothing noisy time series data and identifying underlying trends or patterns. In weather forecasting, time series data represents a sequence of measurements or observations taken at regular

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue VI June 2026

time intervals, such as temperature readings recorded daily or hourly. The Simple Moving Average is a smoothing technique that calculates the average of a fixed number of consecutive data points, often referred to as a "window" or "period." The moving average window moves through the time series, one step at a time, to create a series of smoothed values. To begin using SMA, the initial step involves deciding on the window size, which dictates the number of data points to factor into each average calculation. The choice of window size depends on the desired level of smoothing and the characteristics of the data. For each time point in the time series, the Simple Moving Average is calculated by taking the arithmetic mean of the data points within the moving window. The formula for calculating the SMA for a specific time point t is:

$$\text{SMA} = \frac{A_1 + A_{2+} \ \ldots\ldots + A_n}{n} \qquad (2)$$

where, A_n is the sum of data points, and n is number of total periods.

Proposed System Architecture

The proposed system design is shown in figure 2.

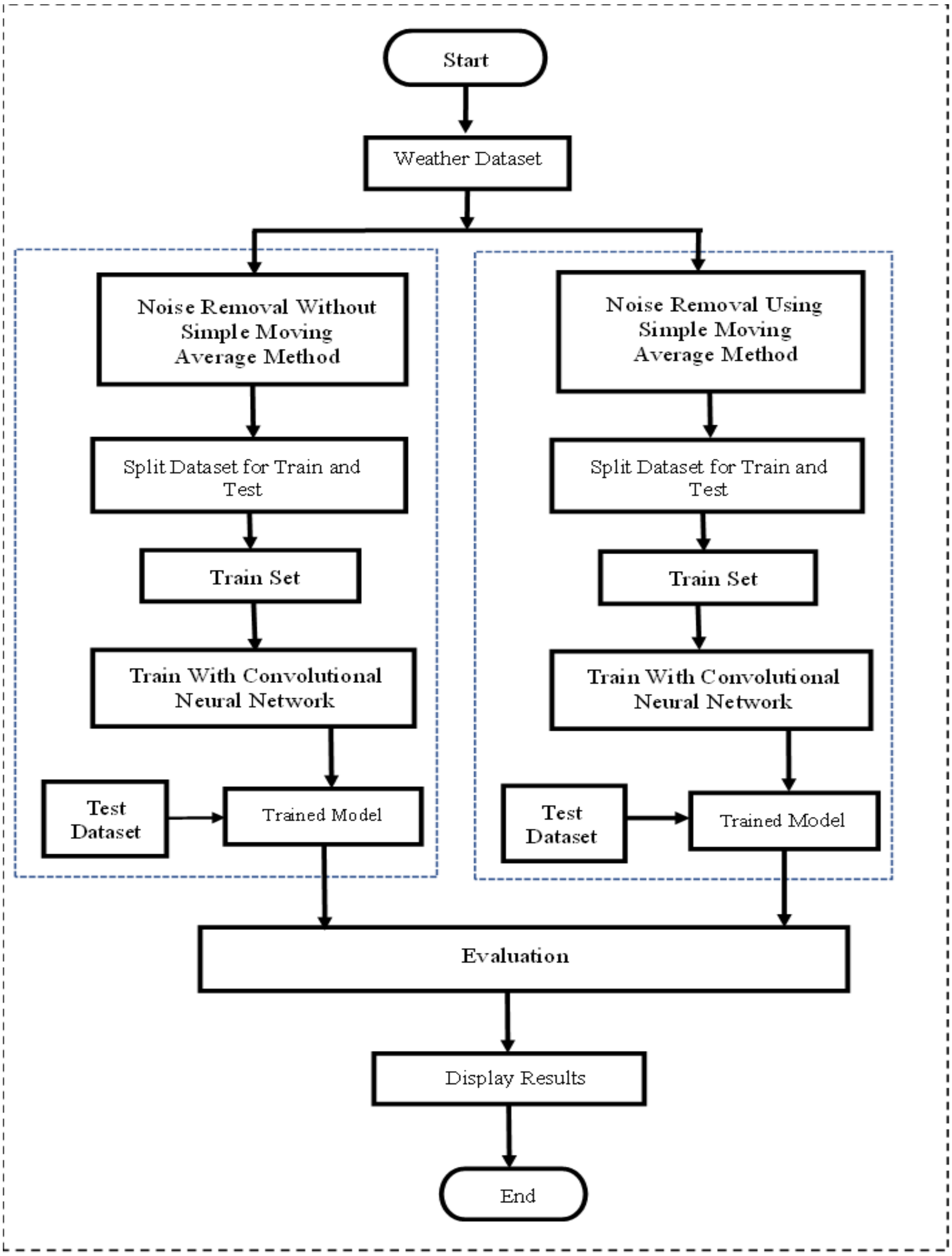

Fig. 2. Proposed System Design

In this figure, ten years history data (from 2013 to 2022) of weather condition in Hmawbi, Myanmar are used to develop the weather prediction model. CNN algorithm is used for developing the weather prediction model. Firstly, we use the original weather data without cleaning the noise in dataset to develop the weather prediction model by using CNN.

Secondly, we use simple moving average method for removing noise in the original dataset and develop the weather prediction model by using CNN and the refined dataset. We compare and evaluate the prediction results of these two developed models.

The original dataset is preprocessed for noise removal with simple moving average. In the preprocessing step, as the collection of data contains useless data and data duplication, these data need to be preprocessed and cleaned for the effective analysis system. The purpose of data preprocessing is for the duplicated and noise data removal as they cannot be used efficiently for predictive analysis. In this step, the prediction task is simplified and cost of processing is significantly decreased in the training stage.

After preprocessing, this dataset is split into training and testing for this system. The training data is input into CNN by tuning hyper-parameters to achieve desired accuracy. The training model is built using CNN with the Adam optimizer is applied. The testing of system is evaluated. Then, this system uses CNN model to predict the weather using the past 3 days and multivariate model. The experiments will be examined for example completeness, appropriateness and quality, and will illustrate the root mean square error and mean absolute error of the used deep learning algorithm. In this system, 32 kernels with 2 convolutional layers are used. For each convolutional layer, maximum pooling and rectified linear unit (ReLU) are used and the pooling size is 4 for only first pooling layer and is 2 for other pooling layers. RELU function is used at fully-connected layer.

Performance Evaluation

In this system, history data of weather condition in Hmawbi, Myanmar is used as input data source. This system model is trained in Kera's which is based on TensorFlow. The system is implemented by the ratio of data size: (Training – 80%, Testing – 20%). Firstly, randomly selected 80% records as training data and 20 % records for testing from this dataset. The key metrics of performance measures (mean squared error and root mean squared error) are evaluated for this proposed system analysis. Mean Squared Error (MSE) is a widely used metric for evaluating the accuracy of a predictive model's forecasts, including time series predictions. It measures the average squared difference between predicted values and actual values over a set of data points. Mathematically, MSE is calculated as shown in formulas for this proposed system analysis.

$$\mathrm{MSE} = \frac{1}{n} * \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \qquad (3)$$

where, n is the total number of data points, y represents the actual (observed) values at each time step, ŷ represents the predicted values at each time step.

Root Mean Squared Error (RMSE) is a variation of MSE and is particularly useful when you want the error metric to be in the same units as the original data. RMSE is the square root of the MSE and is calculated as follows:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} \|y_i - \hat{y}_i\|^2}{n}} \qquad (4)$$

where n is the number of data points, y(i) is the i-th measurement, and ŷ(i) is its corresponding prediction.

Table 1 shows the comparative results for the performance of proposed model for next day prediction, only CNN and SMA with CNN.

Table 1. RMSE and MSE Results

Models	RMSE	MSE
Noise Removal Without SMA	5.398	29.020
Noise Removal With SMA	**3.767**	**14.841**

According to evaluation results, simple moving average with CNN model is considered acceptable for time series data analysis system.

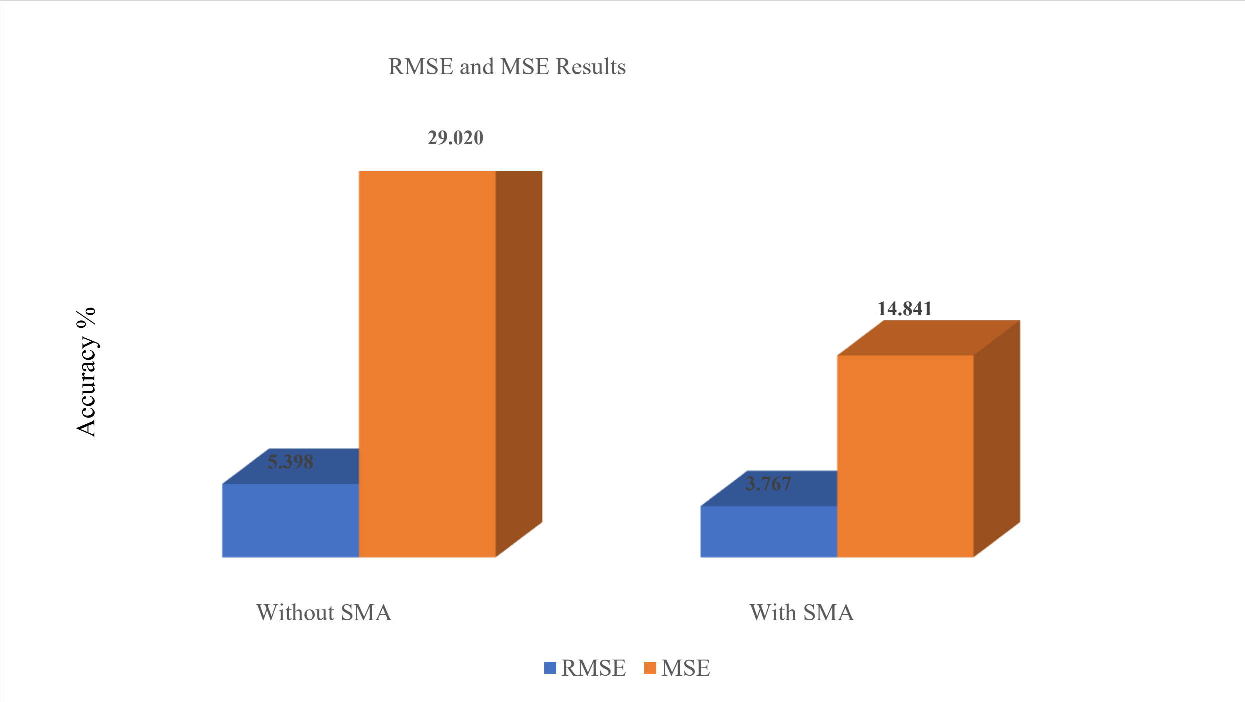

Fig. 3. Analysis Results of Weather Prediction in Difference Models (CNN and CNN with SMA)

According to table1 and figure 3, to enhance the performance of our weather prediction model, we recognized the importance of noise reduction in our dataset. We explored the application of the simple moving average method as a preprocessing step to effectively filter out noise. By implementing the simple moving average method, we successfully improved the quality of our dataset, making it more suitable for training our CNN-based weather prediction model. According to the result of ground truth weather data and the predict weather data is acceptable by using simple moving average method for time series data analysis system.

CONCLUSION

This study utilized CNN and CNN with SMA models to predict daily weather attributes; maximum temperature, minimum temperature in Hmawbi, Yangon, Myanmar. The required weather data was sourced from the Department of Meteorology and Hydrology in Yangon, Myanmar, spanning from 2013 to 2022. The forecasting accuracies of these models were compared using RMSE and MSE metrics. The findings revealed that the CNN model outperformed . CNN with SMA models in terms of both RMSE and MSE due to its superior forecasting accuracy and minimal error.

REFERENCES

1. Guici Chen, Sijia Liu, and Feng Jiang. “Daily Weather Forecasting Based on Deep Learning Model: A Case Study of Shenzhen City, China.” Atmosphere, 13(8), 1208, 2022.
2. Keerthana P. J., Rajeswari P., Amirtharaj R., and PandiyaRajan G. “Weather Forecasting using Deep Learning Algorithm.” International Journal of Special Education, 37(3), 2022.

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue VI June 2026

3. P. S. Mung and S. Phyu. "Time Series Weather Data Forecasting Using Deep Learning." ICCA 2019, 2019.
4. N. Shobha and T. Asha. "Monitoring Weather-Based Meteorological Data: Clustering Approach for Analysis." International Conference on Innovative Mechanisms for Industry Applications (ICIMIA), 2017, pp. 75–81.
5. S. Navadia, P. Yadav, J. Thomas, and S. Shaikh. "Weather Prediction: A Novel Approach for Measuring and Analyzing Weather Data." International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud), 2017, pp. 414–417.
6. V. Leijen, M. Boone, and K. Horgan. "Making Numerical Weather Predictions Portable: Compression of Weather Data for Use in Radar Propagation Modeling." USNC-URSI Radio Science Meeting (Joint with AP-S Symposium), 2017, pp. 12–19.
7. R. Dalmau, M. Perez-Batlle, and X. Prats. "Estimation and Prediction of Weather Variables from Surveillance Data Using Spatio-Temporal Kriging." IEEE/AIAA 36th Digital Avionics Systems Conference (DASC), 2017, pp. 1–8.
8. Y. LeCun, Y. Bengio, and G. Hinton. "Deep Learning." Nature, 521, 2015, pp. 436–444. TechTarget. "Deep Learning (Deep Neural Network) Definition." https://www.techtarget.com/searchenterpriseai/definition/deep-learning-deep-neural-network.