Deep Learning based Weather Forecasting

July Lwin
University of Computer Studies, Yangon
Myanmar
julylwin@ucsy.edu.mm

Tin Zar Thaw
University of Computer Studies, Yangon
Myanmar
tinzarthaw@ucsy.edu.mm

***Abstract*— Weather forecasting is a subject of significant interest and research across multiple scientific communities due to its profound impact on various aspects of daily life worldwide. Over the years, researchers have explored various techniques, including sliding window and different machine learning methods, to improve the accuracy of weather predictions. However, as the demand for more precise and reliable weather forecasts has grown, deep learning techniques have gained popularity for this purpose. This paper aims to predict various weather attributes, including minimum temperature, maximum temperature, humidity, weather type, cloud amount, and wind speed, utilizing a deep learning methodology known as Long Short-Term Memory Network (LSTM). To enhance the accuracy and performance of the weather prediction model, a smoothing algorithm, the Exponential Moving Average (EMA) is being employed to reduce the noise and variation in time series data.**

Keywords*— *Weather Forecasting; Long Short-Term Memory Network; Deep Learning; Exponential Moving Average

I. Introduction

Weather forecasting is a critical application of scientific and technological advancements aimed at predicting atmospheric conditions, including humidity, wind patterns, temperature, and more, for a specific location at future time intervals. The components of weather, such as humidity, wind speed, and temperature, exert a significant impact on diverse aspects of human existence. Computer-based models are popular in use which take the atmosphere information for prediction process. There is still a need for human input to choose the best forecast model, on which the quality of forecast will depend such as pattern recognition, teleconnections, knowledge of model performance and bias.

Machine learning (ML) methods became popular and considering many ML-based methods developed in forecasting applications. But to make intelligence decision, deep learnings are used in this experiment as deep learning structures algorithms in layers to create an “artificial neural network” that can learn and make intelligent decisions on its own. In this paper, a weather forecasting model is developed using a LSTM deep learning approach, aimed at providing highly accurate results tailored to the agricultural sector's needs. Historical weather data from Yangon, Myanmar, is utilized to train the LSTM deep learning model, applying data from the past ten years to mitigate the impact of significant weather variations over the years. The paper's contributions include:

(i) The application of LSTM is taken for weather forecasting.

(ii) The utilization of exponential moving average is applied for noise reduction in time series data, thereby enhancing the accuracy of classification results. The remainder of this paper is structured as follows: Section 2 outlines related works, while Section 3 provides background theory, Section 4 discusses the proposed system, and explain the result and discussion. Finally, in Section 5, conclusions are drawn.

II. Related works

The authors employed [6] a Deep Learning-Based Weather Prediction model utilizing Long Short-Term Memory (LSTM). The quality of the input datasets has a substantial impact on the precision of the forecast outcomes. Sensors and autonomous observation platforms, including those based in the ocean, have accumulated extensive meteorological observation data amounting to petabytes.

In this paper [7], weather attributes such as minimum temperature, maximum temperature, humidity, and wind speed were forecasted using various deep learning techniques: convolutional neural networks (CNN), long short-term memory (LSTM), and an ensemble approach combining CNN and LSTM (CNN-LSTM). The research aimed to predict one-month weather conditions. The experimental results indicated that the ensemble CNN-LSTM model outperformed the other individual methods.

The author introduced the application of the geostatistical interpolation technique known as Kriging for the short-term forecasting of weather conditions [8]. It has proven effective in accurately representing the spatial-temporal distribution of wind and temperature, thereby enabling the production of high-quality, localized short-term weather predictions, complete with measurements of uncertainty. A thorough evaluation, including cross-validation using multiple methods, was carried out to demonstrate the precision of capturing the spatial-temporal distribution of these weather variables through the generated wind and temperature models.

The scientists collected historical weather data pertaining to the Shenzhen region from a range of sources, encompassing parameters like temperature, humidity, wind speed, precipitation, and atmospheric pressure [2]. Artificial neural network model combines the strengths of convolutional neural networks (CNNs) and recurrent neural networks (RNNs) to capture both spatial and temporal patterns within the weather data. CNNs are particularly adept at identifying spatial patterns, such as variations in temperature across different geographical locations, while RNNs excel in capturing temporal patterns, including daily and seasonal weather fluctuations.

In the paper [5], the authors outlined the utilization of data mining techniques to analyze agricultural meteorological data collected from the meteorological center in Bengaluru district. The analysis involved clustering different crop types, including mangoes, grapes, potatoes, and others with hierarchical clustering and K-means clustering, which play a crucial role in decision-making processes aimed at ensuring sustainable agriculture and achieving accurate predictive outcomes. The experimental findings indicated that the hierarchical clustering approach outperformed K-means clustering in terms of its effectiveness in predicting weather conditions.

The authors introduced a system designed to enhance the accuracy of rainfall forecasting through the application of big data predictive analytics within the Hadoop framework [9]. This model aimed to capture relationships among numerous data factors, allowing for the establishment of scoring or weighting patterns for rainfall prediction based on historical data. The system demonstrated its efficacy in processing vast amounts of data through the utilization of big data analytics techniques. Within this system, analytics were conducted through the classification of weather types using the Naive Bayes algorithm, with a specific focus on the humidity weather attribute.

III. BACKGROUND THEORY

The Exponential Moving Average (EMA) is an effective smoothing algorithm employed for the purpose of reducing noise and fluctuations within time series data. It assigns greater significance to recent data points while diminishing the influence of older ones [1]. The EMA at time t, designated as $EMA(t)$, is calculated as illustrated in equation (1) below:

$$(t) = \alpha * X(t) + (1-\alpha) * EMA(t-1) \quad (1)$$

In this equation, $X(t)$ represents the input value at time t and α denotes the smoothing factor (between 0 and 1).

The Long Short-Term Memory (LSTM) is a specific type of recurrent neural network (RNN) designed to capture causal relationships over extended time periods. Within LSTM, gate mechanisms are employed to regulate the flow of data in and out of memory cells, enabling long-term data storage. The model's parameters are detailed in equation (2) below:

$$i_t = (W_i * [h_{t-1}, x_t] + b_i)$$
$$f_t = (W_f * [h_{t-1}, x_t] + b_f)$$
$$o_t = (W_o * [h_{t-1}, x_t] + b_o)$$
$$c_t = f_t * c_{t-1} + i_t * tanh(W_c * [h_{t-1}, x_t] + b_c)$$
$$h_t = o_t * tanh(c_t) \quad (2)$$

In equation (2), i_t, f_t, and o_t correspond to the input, forget, and output gates, respectively, while c_t and h_t represent the cell state and hidden state, respectively. Additionally, W_i, W_f, W_o, and W_c are the weight matrices, and b_i, b_f, b_o, and b_c are the bias terms. Finally, x_t denotes the input at time t.

EMA based LSTM is an enhanced version of the LSTM model that integrates the Exponential Moving Average (EMA) smoothing algorithm to enhance the accuracy of multi-step vector output predictions in time series analysis. The primary distinction between the conventional LSTM and EMA based LSTM lies in how the output is refined within the EMA based LSTM framework. In the standard LSTM, the model generates output directly based on the hidden state of the LSTM cells. Conversely, in EMA based LSTM, the output undergoes smoothing through the EMA algorithm. EMA considers previous predictions and employs a weighted scheme that assigns greater significance to recent predictions, resulting in a more streamlined output. The equations governing EMA based LSTM are outlined in (3) as follows:

$$i_t = (W_i * [h_{t-1}, EMA(x_t)] + b_i)$$
$$f_t = (W_f * [h_{t-1}, EMA(xt)] + b_f)$$
$$ot = (Wo * [ht-1, EMA(xt)] + bo)$$
$$c_t = f_t * c_{t-1} + i_t * tanh(W_c * [h_{t-1}, EMA(xt)] + b_c)$$
$$ht = ot * tanh(ct) \quad (3)$$

Here, (xt) represents the Exponential Moving Average of the input at time t [4].

IV. PROPOSED SYSTEM

The system design is described in Figure 1. Initially, the dataset is partitioned into training and testing subsets for the system. Noise removal in time series weather forecasting systems is a crucial preprocessing step aimed at enhancing the accuracy and reliability of weather predictions. Noise in weather data can arise from various sources, including sensor inaccuracies, measurement errors, and environmental factors. In the preprocessing step, the original dataset is preprocessed to eliminate noise with exponential moving average. The proposed system trains the training dataset with two deep learning methods: LSTM networks [10] and EMA based LSTM networks [4,11] for the weather information of the years 2010 to 2021 in the region of Yangon in Myanmar. This weather dataset consists of 6 attributes, such as date, daily maximum temperature, daily minimum temperature, daily relative humidity, wind speed, and cloud amount. Figure 2 presents original weather dataset.

The system's testing phase is conducted to assess its performance. Subsequently, this system employs the LSTM model and the EMA based LSTM model for weather information of year 2022. The experiments will be scrutinized for comprehensiveness, suitability, and quality, with a focus on evaluating the root mean square error (RMSE) and mean square error (MSE) associated with the deep learning algorithm employed.

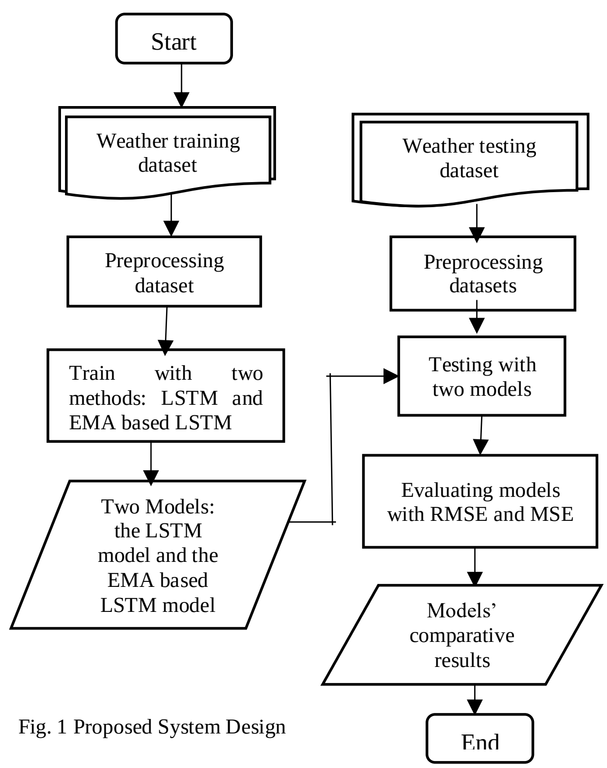

Fig. 1 Proposed System Design

Date	Daily Maximum Temperature (°C)	Daily Minimum Temperature (°C)	Daily Relative Humidity	Wind Speed	Cloud Amount
1/1/2013	32.5	17.0	90	0.0	5
2/1/2013	32.0	17.0	90	0.0	4
3/1/2013	32.0	18.0	91	0.0	5
4/1/2013	32.0	17.0	90	0.0	5
5/1/2013	32.5	18.0	90	0.0	5

Fig. 2 Weather Dataset

V. RESULT AND DISCUSSION

In this system, historical weather data from Yangon, Myanmar serves as the input data source. The system's model is trained using Keras, which is based on the Tensorflow framework. This system aims to train LSTM model with various monthly weather attributes, including minimum temperature, maximum temperature, humidity, weather type, cloud amount, and wind speed from the years 2010 to 2021 with LSTM for Yangon, Myanmar. And then the proposed system predicts the weather information of year 2022 using the trained LSTM model and EMA based LSTM model. In this section, the LSTM sample model is trained with only the monthly average temperature (maximum) over Kaba Aye station during the period of 2001 to 2009 in the figure 3 and predicted the monthly average temperatures (maximum) of 2010 [3].

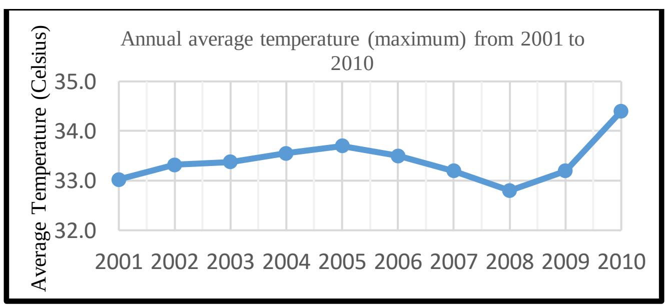

Fig. 3 Annual average temperature (maximum) information over Kaba Aye station.

The system's performance is assessed using key metrics, including root mean square error (RMSE) and mean square error (MSE). According to the sample evaluation result, the system can predict the monthly average temperatures (maximum) of 2010 with RMSE of 1.66 and MSE of 2.64. The proposed system intended a comparative analysis of only LSTM and EMA based LSTM models' performances for 2022 monthly weather prediction with evaluation metrics: RMSE and MSE to choose the best one model.

VI. CONCLUSION

Weather forecasting holds significant importance in the context of everyday human life. In pursuit of enhanced accuracy, deep learning methods are employed for weather prediction, surpassing the performance of traditional machine learning techniques in this experimental setup. This system aims to forecast weather data for the Yangon Region, utilizing LSTM network and EMA based LSTM network to achieve the desired prediction accuracy. Based on the experiment result, the proposed LSTM model can predict the weather information of year 2010 with RMSE of 1.66 and MSE of 2.64. To improve the accuracy of the weather prediction model, data preprocessing techniques are needed to consider and assemble to LSTM model for mitigating the effects of noise in the data. For future research, the proposed system will create two models only LSTM and EMA based LSTM models for the weather information of the years 2010 to 2021 in the region of Yangon in Myanmar and then will analysis these two models with evaluation metrics: RMAE and MSE to choose the best one model.

REFERENCES

[1] D. Haynes, S. Corns and G. K. Venayagamoorthy, "An Exponential Moving Average algorithm," *2012 IEEE Congress on Evolutionary Computation*, Brisbane, QLD, Australia, 2012.

[2] G. Chen , S. Liu, and F. Jiang, "Daily Weather Forecasting Based on Deep Learning Model:A Case Study of Shenzhen City, China", Special Issue Industrial Air Pollution Control in China, Atmosphere 2022, 13(8), 1208.

[3] K. H. Paing, Y. Li, Z. M. M. Sein. "Temperature (maximum) and Rainfall Prediction Using FNNand RNN Models over Kaba Aye Station, Yangon Region, Myanmar", North American Academic Research, 6(3), 1-13, 2023.

[4] M. Diqi, A. Sahal, F. N. Aini, "Multi-Step Vector Output Prediction of Time Series Using EMA LSTM", JOIN (Jurnal Online Informatika) p-ISSN: 2528-1682, e-ISSN: 2527-9165 Volume 8.

[5] N. Shobha and T. Asha, "Monitoring weather based meteorological data: Clustering approach for analysis", the International Conference on Innovative Mechanisms for Industry Applications (ICIMIA), 2017.

[6] P. J. Keerthana, P. Rajeswari., Amirtharaj. R, PandiyaRajan. G, "Weather Forecasting using Deep Learning Algorithm", International Journal of Special Education, Vol.37, No.3, 2022.

[7] P. S. Mung, and S. Phyu, "Time Series Weather Data Forecasting Using Deep Learning", ICCA 2019, 2019.

[8] 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.

[9] S. Navadia, P. Yadav, J. Thomas and S. Shaikh, "Weather prediction: A novel approach for measuring and analyzing weather data", the International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), 2017, pp.414-417.

[10] Y. LeCun, Y. Bengio, and G. Hinton, "Deep Learning", vol. 521, Nature, London, 27 May, 2015, pp. 436-444.

[11] https://corporatefinanceinstitute.com/resources/equities/exponential-moving-average-wma/.