Time Series Weather Data Forecasting Using Deep Learning

Pau Suan Mung
University of Computer Studies, Yangon
Yangon, Myanmar
pausuanmung@ucsy.edu.mm

Sabai Phyu
University of Computer Studies, Yangon
Yangon, Myanmar
sabaiphyu@ucsy.edu.mm

***Abstract*— Weather forecasting is an interesting research in a number of applications and has great attention of researchers from various research communities due to its effect on the daily life of human globally. Through the years, researchers used sliding window and different machine learning techniques for this purpose. Weather data is in the form of time series data. Time series data can be forecasted using regression. For better accuracy and performance, deep learning becomes popular for prediction. The aim of this research work is to predict weather attributes such as minimum temperature, maximum temperature, humidity and wind speed using different deep learning methods, namely convolutional neural networks (CNN), long short-term memory (LSTM) and ensemble of CNN and LSTM (CNN-LSTM). Dataset including daily weather conditions from past two decades of Yangon Region, Myanmar is 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 shows that ensemble CNN-LSTM model has better performance than other individual methods.**

Keywords—Deep Learning, Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), Root Mean Square Errors (RMSE)

I. Introduction

Weather forecasting is the prediction application in science and technology using the conditions of atmosphere in a given location and time. It is important in the process of air traffic, marine, industry, agriculture, forestry and even public health sector as some diseases such as dengue and malaria are closely related to the influence of weather condition. Actually, it effects on our daily life.

Along the years, people have attempted to forecast the weather informally and started the formal forecasting in the 19th century. Based on the collected quantitative data of the current state of atmosphere in a given place and time, this forecasting can be made for such location and can also learn how the weather changes using some meteorology projects. Some information taken for this prediction are barometric pressure, current weather conditions, sky condition and cloud cover. Computer-based models are popular in use which take the atmosphere information for prediction process. [1] 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. Sometime the forecast is inaccurate because of the unstable nature of atmosphere, the requirement of massive computational power for solving complex equations that implement the situation of the atmosphere, the error occurred in the initial condition measurement, inadequate understanding of the process of atmosphere. In some cases, the forecast may become less accurate due to the different between the situation of current time and the time for forecasting to be made. The use of ensemble methods and model consensus can narrow the error and will help pick up the most suitable model and more accurate forecast.

There is a large amount of end users to weather forecast. Based on prediction, weather warnings are made and these warnings can be used to protect life and property. Forecasting on temperature and air precipitation are vital to agriculture and are used by companies to estimate some demands for coming days. Weather forecast such as heavy rain, snow and wind chill, is useful to determine the dresses to be worn in a given day. Our daily life is closely related to the weather and sometimes it is even essential for our survival. Weather condition may affect how we feel better or not. It affects can also be noticeable on the food we eat, the business we process and the well-being we face. Therefore, accurate and timely forecasting of weather data is imperative so that the necessary precautions can be taken to minimize weather-associated risks.

Weather forecasting attracts great attention of researchers from various research communities because of its effect to human life globally. Rresearchers explore patterns that are hidden in the large dataset using the current wide availability of massive weather observation data and make weather prediction. The advent of computer technology and data science in the last decades also support the research of weather forecasting to get more accurate results.

Weather data is statistical and some statistical analysis methods were applied on forecasting. Sliding winding is one of the most famous statistical methods. Then machine learning (ML) approaches became popular and considering the large number of ML-based approaches developed in forecasting applications. Machine learning uses algorithms for parsing data, learning from that data, and making informed decisions based on what it has learned. But to make intelligence decision, deep learnings are used in this experiment because deep learning structures algorithms in layers to create an "artificial neural network" that can learn and make intelligent decisions on its own. [2]

Two popular deep learning methods, namely CNN and LSTM are used to predict and the third method, CNN-LSTM, is also used for better performance. History data of weather condition in Yangon, Myanmar is taken from Open Weather [3] organization. Although it includes previous data of over 50 years of weather condition, this experiment uses previous 20 years data to avoid large difference in weather changes along the years.

This paper is organized as follows. The next section presents related research works and the following section includes brief of background theories such as convolutional neural network (CNN), long short-term memory (LSTM) and the combination of CNN and LSTM in deep learning. Section 4 is detail description of experiment and the next part of this section is experimental result. The step-by-step processes in this experiment are included. This paper concludes with Section 5, which includes conclusion and further extension of this research work.

II. Related Works

Here are some related research works on forecasting weather condition.

In the paper [4], the authors proposed weather forecasting for flight navigation area in which visibility is important factor for all stages of flight such as during taxi-out, take-off, initial climb, landing and taxi-in. In this paper, the authors took weather forecasting with Autoregressive Integrated Moving Average (ARIMA) and Long Short-Term Memory (LSTM) model. Weather data such as visibility, temperature, dew point and humidity from Hang Nadim Airport, Batam Indonesia, were collected and these were in the form of time series data. Such 40026 time series weather data were used to predict with these two methods and compared their performance. Visibility is the most important weather variable and therefore it had big impact big impact on all phases of flight, especially close to the ground and another weather data such as temperature, humidity and dew point variable were used as intermediate variable. Root Mean Square Error (RMSE) was used for comparing these models. The paper showed that LSTM model is better than ARIMA model.

The authors in the paper [5], proposed a new temperature prediction model using deep learning on real observed weather data. But the collected data was incomplete and extended gaps because there was a limitation in collection weather data. Long Short-Term memory (LSTM), a kind of recurrent neural network was used in this experiment because the temperature data is seasonal and time-series. Different investigation on LSTM configurations were made and the proposed LSTM-based model could reflect the time-series nature of temperature data. Data was restored using proposed model's refinement function for missing a part of data. The LSTM model was retained when all missing data was refined. This system could predict the temperature for 7 to 14 days with 6, 12 and 24 hours. For accuracy analysis, the Root Mean Square Error (RMSE) of refined and proposed LSTM model was compared to those of a mathematical model currently used in the meteorological office in South Korea. This experiment showed that the RMSE of refined LSTM is the lowest for temperature prediction in other prediction methods. And the prediction accuracy of the proposed model was higher than those of other models. The dataset for this experiment was taken from the Meteorological Office of South Korea (KMA) for 37 years from 1981 to 2017. The temperature was predicted for August 2017 to October 2017.

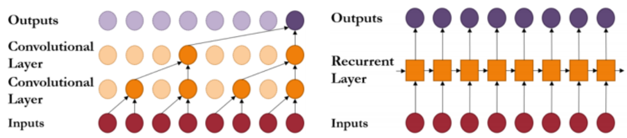
One of the papers [6] proposed a weather prediction system to be used in the agriculture sector in India because there are many cultivation fields around the country. The weather attributes such as minimum temperature, maximum temperature, precipitation rate., wind speed and temperature are many important roles in the field of cultivation. Weather prediction technology allows farmers for a chance for better productivity of crops. In this system, the authors proposed Artificial Neural Network (ANN), a soft computing technique. One of ANN technique, Back Propagation Algorithm, was used on 3 years (2009-2011) weather data gathered by the Regional Meteorological Center (RMC) of Indian Meteorological Department (IMD). Mean Square Error (MSE) was used to generate the error rate of this proposed system and Matlab programming language was used for developing this system.

Rainfalls are often sudden temporally and short events, spatially varying in size, and economically important. The successful prediction of spatial, temporal details and dynamics of rains is important to societal welfare. The paper [7] proposed a prototype of deep learning for rainfall prediction for Beijing City, China. The rainfall for the city was predicted in a short time as it predicted for the next two hours. This experiment used spatial resolution in 1km and temporal resolution in 6 minutes. The computation time was also less than 10 seconds and thus it was better than traditional weather forecast. The deep learning network was combined with U-Net, ResNet, Sqeeze-and-Excitation and the spatial attention module. The difference between other systems is that this system fully relied on full convolution layer. FCN approach was suggested for better performance of meteorologically assessment metrics. This experiment was the first system to be served by a capital weather service and it is now up and running since the rain season on summer in 2019. Currently, Weather forecast office at Beijing Haidian Meteorological Bureau is using this proposed system.

In the paper [8], the authors proposed weather forecasting to be used in agricultural and industrial sectors because these sectors are principally dependent on weather situations. This system also proposed to use long short-term memory (LSTM) technique and for different combinations, training is taken in neural networks. The parameters used for weather data were temperature, pressure, humidity, dew points, wind speed, precipitation and visibility. The historical dataset on weather forecast was taken from National Climate Data Center (NCDC) from November 2007 to October 2017. The block diagram of this system was presented and the accuracy results were also described.

III. Deep Learning for Time Series Data Forecasting

Deep learning, one of the machine learning techniques, teaches computers to do naturally like humans and to learn by example. Computer models in deep learning learn to perform classification or regression tasks directly from input data such as text, song, image and video. In deep learning, based on a large set of data and neural network, models are trained and used for prediction. Deep learning models can take accuracy with state-of-the-art. [9] Two types of deep learning, CNN and RNN are shown in Fig. 1.



(a) CNN Model. (b) RNN Model.

Fig. 1. Deep learning neural networks

Deep learning can be used for the forecasting of time series data. The model of time series forecasting predicts the future values of a target $y_{i,t}$ for a given entity i at time t. a logical grouping of temporal information is represented by each entity that is vital signs from different patients in medical fields and measurements from each weather stations and they are observed the same time. [10] One-step-ahead forecasting models take the form:

$$\hat{y}_{i,t+1} = f(y_{i,t-k:t}, x_{i,t-k:t}, s_i) \quad (1)$$

where $\hat{y}_{i,t+1}$ is the forecast of the model, $y_{i,t-k:t} = \{y_{i,t-k}, \ldots, y_{i,t}\}$ and $x_{i,t-k:t} = \{x_{i,t-k}, \ldots, x_{i,t}\}$ are target observations and exogenous inputs respectively over a look-back window k, s_i is static metadata associated with the entity (e.g. location of sensor), and $f(.)$ is the prediction function learnt by the model. From the univariate forecasting (i.e. 1-D targets), the same components can be used to extend to multivariate models with no loss of generality. [10]

A. Convolutional Neural Networks (CNN)

Convolutional Neural Network (CNN) is designed traditionally for image datasets and can collect local relationship that are invariant across spatial dimensions. But to adapt the CNN network in time series dataset, multiple layers of causal convolutions are used. For forecasting purpose, convolutional filters are designed to ensure only past information. For the intermediate feature at hidden layer l, each causal convolutional filter is taken as in the following:

$$h_t^{l+1} = A\left((W * h)(l, t)\right) \quad (2)$$

$$((W * h)(l, t) = \sum_{i=0}^{k} W(l, i)\, h^l_{t-i} \quad (3)$$

Where h^l_t is an intermediate state at layer l at time t, and $*$ is the convolution operator. $W(l, \tau)$ is a fixed filter weight at layer l, and $A(.)$ is an activation function, or a sigmoid function, that represents any non-linear processing of architecture-specific. For CNNs that use a total of L convolutional layers, the encoder output is then $z_t = h_t^L$. [10]

B. Long Short-Term Memory (LSTM)

LSTM (Long-Short Term Memory network) is a recurrent neural network that allows for the accounting of sequential dependencies in a time series. The older variants of RNN can suffer from limitations for learning long-range of data dependencies due to the infinite lookback window and due to the issues of exploding and vanishing gradients. These limitations can be seen as a form of resonance in the memory state. To address these limitations, LSTM model was developed. By improving gradient flow with the network, LSTM is achieved from the use of a cell state c_t that stores long-term information. [10] LSTM modulated a series of gates as follows:

$$\text{Input gate:} \quad i_t = \sigma(W_{i1}z_{t-1} + W_{i2}y_t + W_{i3}x_t + W_{i4}s + b_i) \quad (4)$$

$$\text{Output gate:} \, o_t = \sigma(W_{O1}z_{t-1} + W_{O2}y_t + W_{O3}x_t + W_{O4}s + b_O) \quad (5)$$

$$\text{Forget gate:} \, f_t = \sigma(W_{f1}z_{t-1} + W_{f2}y_t + W_{f3}x_t + W_{f4}s + b_f) \quad (6)$$

Where z_{t-1} is the hidden state of LSTM, and $\sigma(.)$ is the sigmoid activation function. The gates that modify the hidden states and cell states of LSTM can be defined as:

$$\text{Hidden state:} \, z_t = O_t \; \Theta \; \tanh(c_t) \quad (7)$$

$$\text{Cell state:} \quad c_t = f_t \; \Theta \, c_{t-1}$$
$$+ \, i_t \; \Theta \tanh\left(W_{C1}z_{t-1} + W_{C2}y_t + W_{C3}x_t + W_{C4}s + b_C\right) \quad (8)$$

Where Θ is the element-wise or Hadamard product, and $tanh(.)$ is the tanh activation function.

C. CNN-LSTM Model

For a given time series, there are correlations existed between observations and this correlation is known as autocorrelation. All such observations are treated as independent in a standard neural network. This may be erroneous and misleading results may be generated.

CNN can be used for a process that is known as convolution which determines the relationship between two functions. The shape of one function modified by other can be easily expressed with the convolution integral. These networks are useful for image classification. And they do not account for sequential dependencies in the way that the recurrent network can do.

However, the main advantage of CNN is dilated convolutions that is suitable for time series forecasting. This is also the ability of using filters to determine dilations between each cell. The neural network can better understand the relationships between different observations in the time series by using the size of the space between each cell. This may cause CNN and LSTM to be combined for time series forecasting. In such combination, LSTM layer can be used to account for sequential dependencies in time series, whereas CNN layer can inform the process through the use of dilated convolutions. The combination of CNN and LSTM is shown in Fig. 2 and in which, CNN network learns input features and an LSTM interprets them in the ensemble of CNN and LSTM (CNN-LSTM).

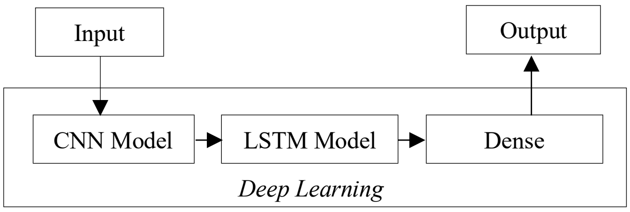


Fig. 2. Workflow of CNN-LSTM

In time series forecasting, the usage of standalone CNN is increasing. There are some impressive results that are produced by the combination of several Conv1D layers. These results can be rivalling with that of a model using CNN and LSTM layers.

D. Root Mean Square Error (RMSE)

Root Mean Square Error (RMSE) is used for finding the prediction error and it is a standard way for measuring error of a model in quantitative data prediction. It is also the value of the difference between predicted value by model and the actual value observed from the environment that is used for modelling. It is commonly used in forecasting, climatology and regression analysis for experimental results. RMSE can tell that how the data is concentrated around the best fit line.

The formula of RMSE is:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(X_{predict,i} - X_{model,i})^2}{n}} \quad (9)$$

where, $X_{predict}$ is predicted values and X_{model} is modelled values at i, and n is the number of samples.

IV. SYSTEM EXPERIMENT

Weather forecasting plays an important role in different types of application. There are many weather attributes. Among them, minimum temperature, maximum temperature, humidity and wind speed are used as its experimental attributes of this finding. Using such attributes, this system intends to investigate the best deep learning method for weather forecasting using univariate time series data. Some famous deep learning algorithms, CNN and LSTM, are compared with the ensemble algorithm of CNN-LSTM and this experiment shows that CNN-LSTM is the best method for weather forecasting than CNN and LSTM algorithm alone. The followings are the step-by-step processes of this experiment.

a. Experimental Setup
b. Data Description and Preprocessing
c. Process of Experiment
d. Experimental Result
e. Performance Analysis

A. Experimental Setup

Tensorflow, an open-source software library for machine learning and deep learning, is free and powerful. Tensorflow can be used in a number of tasks and it is used particularly on training and inference of deep neural networks. This experiment is carried out with Tensorflow on Python programming language. The PyCham IDE is set up with the following details:

- Python 3.6.0
- Tensorflow 1.13.1
- Pandas 0.24.1
- Numpy 1.16.0
- Keras 2.2.4

B. Data Description and Preprocessing

The weather dataset of Yangon region, Myanmar is taken from Open Weather web site [3] and minimum temperature, maximum temperature, humidity and wind speed are used as experimental attributes. These weather attributes, their units of measurement and data types are shown in Table I.

TABLE I. DATA DESCRIPTION FOR WEATHER ATTRIBUTES

Weather Attributes	Unit of Measure	Data Type
Minimum Temperature	Degree Celsius (°C)	Numerical
Maximum Temperature	Degree Celsius (°C)	Numerical
Humidity	Percentage (%)	Numerical
Wind Speed	Miles per Hour (mph)	Numerical

In this experiment, history data for previous 20 years (January 2000 to April 2020) is used as time series data. The sample of dataset is shown in Table II. This dataset includes 99,360 data instances or weather conditions hourly and the instances are averaged on the daily basis for getting daily measures because this experiment intends to predict daily weather data for one month, May 2020. These average values are saved as csv file format to be used by Tensorflow in Python. In this table, +0630 UTC indicates the country Myanmar.

TABLE II. SAMPLE OF DATASET

Date/Time	Minimum Temperature (°C)	Maximum Temperature (°C)	Humidity (%)	Wind Speed (mph)
2000-01-01 00:00:00 +0630 UTC	10.00	14.46	86	1.34
2000-01-01 01:00:00 +0630 UTC	10.00	13.63	86	0.62
2000-01-01 02:00:00 +0630 UTC	12.82	16.50	87	0.23
…	…	…	…	…
…	…	…	…	…
2020-04-30 23:00:00 +0630 UTC	20.48	25.56	52	2.1

The dataset with daily record is divided into 4 univariate time series datasets and each of which includes a pair of day and value for each weather attribute. After data preprocessing step, 4,140 instances of each weather attribute are ready for further processing.

C. Process of Experiment

After preprocessing step, weather data is collected on daily basis for 20 years and this dataset includes day and value pair for each day. This dataset is taken as the training dataset and input for deep learning algorithms such as CNN, LSTM and CNN-LSTM. The framework of this experiment is shown in Fig. 3.

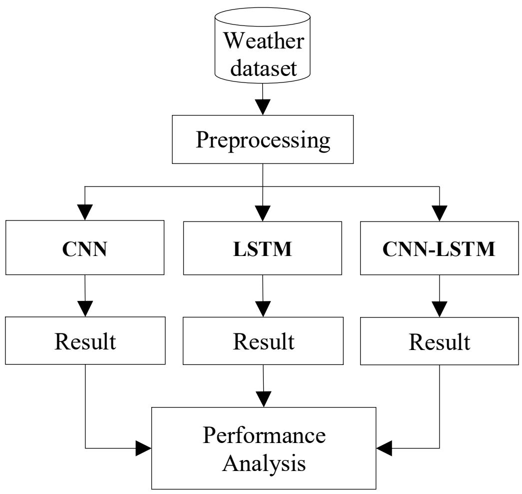


Fig. 3. Framework of experiment

The system builds the model for each weather attribute by using respective training dataset of 20-year weather conditions and these models are tested for 31 days of May 2020. Each algorithm is run for 30 times and the results are averaged and recorded as its output. The process continues until the outputs for all weather attributes are produced. All the outputs are collected and compared for performance analysis.

D. Experimental Result

After completing all processes, the system produces its prediction for each algorithm on each weather attribute for all days in May 2020. The sample of this prediction output is shown in Table III, in which actual and predicted values are included.

TABLE III. SAMPLE OF ACTUAL AND PREDICTED VALUES

Date	Actual/ Algorithm Name	Min Temp (°C)	Max Temp (°C)	Humidity (%)	Wind Speed (mph)
2020-05-01	***Actual***	***22.73***	***30.96***	***60.08***	***2.29***
	CNN	21.55	30.21	67.41	2.47
	LSTM	22.26	30.08	75.55	2.09
	CNN-LSTM	22.02	30.12	72.11	2.15
2020-05-02	***Actual***	***22.75***	***31.26***	***69.46***	***2.44***
	CNN	22.81	31.19	58.11	2.94
	LSTM	23.77	32.84	43.63	2.65
	CNN-LSTM	22.06	31.22	60.04	3.14
2020-05-03	***Actual***	***24.06***	***32.65***	***55.79***	***2.31***
	CNN	22.18	31.57	76.19	2.45
	LSTM	23.40	33.04	73.91	2.50
	CNN-LSTM	22.41	31.29	64.34	2.55
…		…	…	…	…
…		…	…	…	…
2020-05-31	***Actual***	***22.58***	***30.85***	***71.42***	***2.42***
	CNN	22.58	32.69	64.75	2.63
	LSTM	22.92	32.27	65.56	2.49
	CNN-LSTM	23.29	32.15	68.10	2.29

The comparisons for prediction of each weather attributes using CNN, LSTM and CNN-LSTM and actual measures are depicted from the Fig. 4 to Fig.7. In each of these figures, the X-axis represents the prediction period, days in May 2020, and Y-axis represents measurement unit for each of weather attributes such as Degree Celsius for Minimum Temperature and Maximum Temperature, Percentage for Humidity and Miles per Hour for Wind Speed.

In Fig. 4 and Fig. 5, the prediction values of Degree Celsius of minimum and maximum temperature are compared. The predictions of humidity percentage are compared in Fig. 6 and the values of predictions on miles per hour (mph) of wind speed are also compared in Fig. 7. As shown in each figure, our predictions are close to the actual measures.

E. Performance Analysis

The Root Mean Square Error or RMSE values from each deep learning technique for each weather attribute are calculated using RMSE equation shown in Eq. 9. Then each RMSE value is recorded for analysis. As the RMSE values are different because of stochastic nature of machine learning algorithm and usage of training data randomness during learning, the average RMSE of the 30 repeated evaluations of the model is recorded. The comparison of RMSE from each method is also shown in the Table IV. Although the RMSE values of CNN are smaller than those of LSTM for all attributes, the combination of CNN-LSTM algorithm is optimal because it gets the smallest RMSE values. It means that CNN-LSTM is the best method for time series univariate data forecasting.

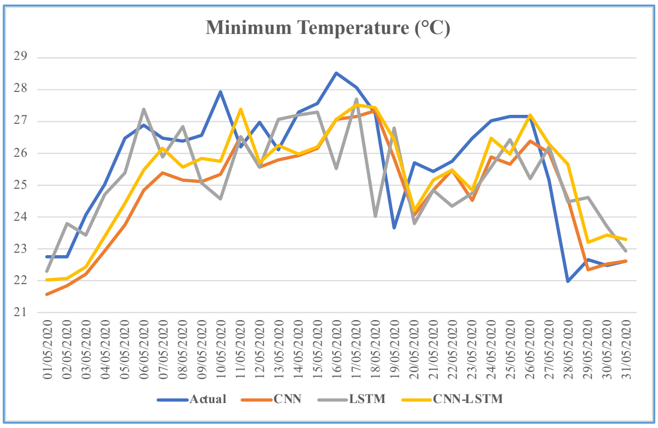


Fig. 4. Comparison of prediction for minimum temperature

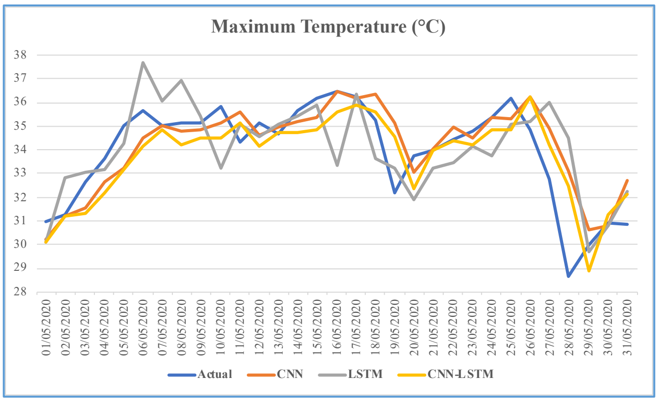


Fig. 5. Comparison of prediction for maximum temperature

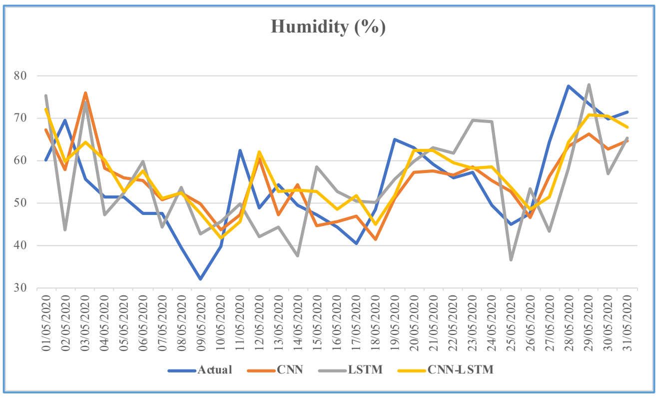


Fig. 6. Comparison of prediction for humidity

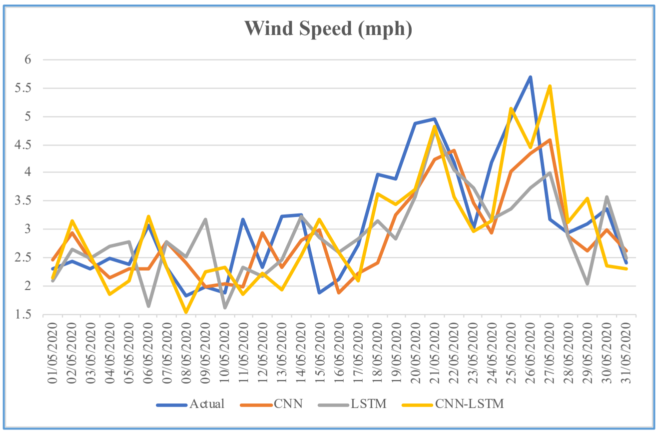


Fig. 7. Comparison of prediction for wind speed

TABLE IV. COMPARISON OF RMSES

Weather Attributes	CNN	LSTM	CNN-LSTM
Minimum Temperature	1.453	1.872	1.370
Maximum Temperature	1.305	1.880	1.283
Humidity	9.025	12.483	8.604
Wind Speed	0.752	0.812	0.702

V. CONCLUSION AND FURTHER EXTENSION

Weather forecasting is very important for human daily life. For weather prediction, deep learning methods are applied for better accuracy than any machine learning methods in this experiment. Weather history data for Yangon, Myanmar is taken from Open Weather and it is preprocessed to be ready for deep learning algorithms. The weather dataset is extracted for 20 years from January 2000 to April 2020 period for avoiding large difference in weather changes, and predicts for May 2020. Different types of neural networks such as convolutional neural network, recurrent neural network and the ensemble of these two networks, are used in this system. The algorithms used are Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM) and CNN-LSTM. For performance analysis, Root Mean Square Error (RMSE) values are recorded and analysis process is taken on such values. According to this study, CNN-LSTM method is the best because its RMSEs are smaller than those of other individual methods.

The main aim of prediction algorithms is for getting more accurate result. In this experiment, three algorithms are used with fixed parameters. The parameters such as the number of hidden layers or epochs can be adjusted for better accuracies. Another extension is to set the training data with suitable data size to get the best model. Weather prediction for one month is made in this experiment but the prediction period can be flexible as needed such as prediction for daily, weekly, monthly or yearly.

REFERENCES

[1] Dirmeyer, Paul A., Schlosser, C. Adam, Brubaker, Kaye L., "Precipitation, Recycling, and Land Memory: An Integrated Analysis" (PDF). Journal of Hydrometeorology. 10 (1): 278–288. Bibcode:2009JHyMe. 10..278D. doi:10.1175/2008JHM1016.1. hdl:1721.1/52326.

[2] https://machinelearningmastery.com/what-is-deep-learning/, "What is Deep Learning?". Jason Brownlee, 2019.

[3] https://home.openweathermap.org/ "OpenWeather–Historical weather for any location".

[4] Afan Galih Salman, Yaya Heryadi, Edi Abdurahman and Wayan Suparta, "Weather Forecasting Using Merged Long Short-Term memory Model (LSTM) and Autoregressive Integrated Moving Average (ARIMA) Model", Journal of Computer Science, 2018, DOI: 10.3844/jcssp..2018..930.938.

[5] Inyoung Park, Hyun Soo Kim, Jiwon Lee, Joon Ha Kim, Chul Han Song and Hong Kook Kim, "Temperature Prediction Using the Missing Data Refinement Model Based on a Long Short-Term Memory Neural Network", Atmosphere Journal, 2019, doi:10.3390/atmos10110718.

[6] Govind Kumar Rahul, Saumya Singh, Saumya Dubey, "Weather Forecasting Using Artificial Neural Networks", 2020 8th International Conference onReliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), 2020 IEEE.

[7] Kuan Song, Gouwei Yang, Qixun Wang, Chunmeng Xu, Gong Zhand, Xiaochen Yu, Zhu Gu, Wenpeng Zhang, "Deep Learning Prediction of Incoming Rainfalls: Operational Service for the City of Beijing China", 2012 International Conference on Data Mining Workshops (ICDMW), 2019 IEEE, DOI: 10.1109/ICDMW.2019.00036.

[8] Dires Negash Fente and Dheeraj Kumar Singh, "Weather Forecasting Using Artificial Neural Network", Proceedings of the 2nd International Conference on Inventive Communication and Computational Technologies (ICICCT 2018), 2018 IEEE.

[9] https://www.mathworks.com/solutions/deep-learning.html?s_tid=hp_brand_deeplearning/, MATLAB for Deep Learning, Data preparation, design, simulation, and deployment for deep neural networks.

[10] Bryan Lim and Stefan Zohern, "Time Series Forecasting with Deep Learning: A Survey", Department of Engineering Science, University of Oxford, UK, The Royal Society Publishing, September 2020.