International Journal of Advanced Engineering and Technology
www.allengineeringjournal.com
ISSN: 2456-7655
Received: 21-12-2022, Accepted: 07-01-2023, Published: 23-01-2023
Volume 7, Issue 1, 2023, Page No. 5-10

Assessment of the potential of various types of long short-term memory (LSTM) artificial neural networks and its application in weather forecasting

Manar Ahmed Mohammed Alblooshi, Fatmah Rashed Mohamed Obaid Alhefeiti, Muhammed Sirajul Huda Kalathingal, Shaher Bano Mirza*, Fouad Lamghari Ridouane
Fujairah Research Centre, Fujairah, United Arab Emirates

Abstract
Predicting the weather accurately is essential for daily life. The systematic recording of meteorological conditions is beneficial to many fields, but particularly to agriculture and other industries that rely on it. The overarching goals of this research are to develop an accurate and flexible statistical model for making city-level weather predictions and assess how intermediate meteorological variables affect the LSTM model's performance. Success with this method on many different challenging prediction issues can be attributed to the sophistication of modern neural networks. The LSTM model's results for weather forecasting are plotted using the standard Python library matplotlib. pyplot. Artificial neural networks with deep layers can model complex data structures. In order to control the data stored in a cell state, Long Short-Term Memory (LSTM) units implement input, output, and forget gates. The error rates of LSTM models are lower than those of other models, allowing them to be used more frequently for predicting. In our analysis, stacked LSTM outperformed both single-cell and bidirectional models when making predictions. The mean square error with single-cell models has an average value of 0.52 based on the weather data we have. The bidirectional approach achieved an astounding 0.99 with the lowest error rate of 0.3 for temperature accuracy, and similarly impressive results were achieved for wind speed accuracy. However, it recorded a low accuracy rating of 0.84 with 31.23 MSE while trying to gauge humidity.

Keywords: artificial neural networks, deep learning, long-short term memory, recurrent neural networks, stacked, single cell, bidirectional, python, weather forecasting, emirates of fujairah, united arab emirates

Introduction

Utilizing physical concepts to anticipate the weather is known as weather prediction. In addition to forecasting meteorological occurrences, it also predicts alterations to the Earth's surface, such as snow and ice cover, storm surges, and floods. Accurate weather forecasts are essential for planning everyday activities. Numerous fields benefit from systematic weather records, particularly agriculture and other industries that rely on it. If it is possible to foresee long-term weather trends, agricultural activities such as planting, and harvesting can be planned and executed more efficiently. However, the weather has a substantial impact on the airline sector. They must be informed of the local weather conditions in order to arrange flights. Each day, they use weather forecasts to make the most informed decisions possible, which may contribute to our safety. In the past decade, research has been conducted on weather forecasting utilizing neural network or NN [1], recurrent neural networks or RNN [2], Temporal Convolutional Networks [3], Artificial Neural Network Model or ANN [4], HWNN model [5], Fuzzy time series model for temperature prediction [6], and NN for short term wind forecasting using LIDAR Data [7]. This project aims to construct a dependable and adaptable statistical model for predicting weather variables in a city, as well as to evaluate the effect of intermediate meteorological variables on the LSTM model's capacity to reliably predict weather conditions (LSTM).

Long short-term memory (LSTM) manages complex subjects in deep learning. It refers to algorithms that seek to mimic human brain activity in order to analyze connections in sequential data. The LSTM deep learning architecture can effortlessly store and imitate the data sequence.

The architecture of the cell in LSTM unit consists of input gates, forget gates, and output gates, as well as four gating levels. Collaborating input gates decide which input is contributed to the cell state. Based on the present state of the cell, the forget gate determines which former state of the cell to erase. The output gates determine which output is delivered via them [8].

Different lstm models

stacked LSTM

The stacked LSTM is an extension of this model consisting of multiple hidden LSTM layers with multiple memory cells per layer. The stacked LSTM hidden layers increase the model's depth, more accurately qualifying it as a deep learning method. The approach's success on a variety of difficult prediction problems is attributed to the depth of neural networks. The stacked LSTM is currently a reliable method for tackling difficult sequence prediction issues. Multiple LSTM layers constitute a stacked LSTM architecture in an LSTM model. The LSTM layer above outputs a sequence rather than a single value to the LSTM layer below. To be more specific, one output for each input time step equals one output time step for the entirety of the input time steps. Therefore, the stacked LSTM was selected for this study [9].

Single cell LSTM model

Each BC calculates the LSTM cell and its dense decoder. Through repeated use, they are able to accumulate long-range knowledge by analyzing the current signal amplitude and output of the previous cell. The network is intended to forecast the transverse energy deposit with a lag of six BCs [10].

Bidirectionnel LSTM

Bidirectional LSTM (BiLSTM) is a recurrent neural network that is mostly utilized for natural language processing. Unlike regular LSTM, input travels in both directions, and information from both sides can be utilized. In addition, it is a potent instrument for modelling the sequential dependencies between words and phrases in both directions. In short, BiLSTM adds an additional LSTM layer that reverses the flow of information. In short, it indicates that the input sequence flows in reverse in the additional LSTM layer. The outputs from both LSTM layers are then combined in a variety of methods, such as by averaging, summing, multiplying, or concatenating. RNNs combined. One network receives the input sequence in regular time order, while another receives it in reverse time order. Typically, the outputs of the two networks are combined at each time step [11].

Materials and methods

Data Collection

In order to use the LSTM (Long Short-Term Memory) model, meteorological data for the city emirates of Fujairah, UAE has been collected for a period of five years, beginning in 2017 and continuing through 2022 till September. We collected information on the temperature, as well as the humidity and the wind. Initially, if the data were collected at a number of different periods, it is best to use a daily average when building the model. This will both reduce the amount of data that is required for the model and ensure its accuracy.

LSTM model

Meteorological forecasting results were generated using three model types: single cell, stacked model, and Bidirectional Cells. To build an LSTM model, it is necessary to use Python's core Keras packages. It is essential to perform LSTM prediction with clean input data. Using data conditioning libraries like as SciPy, you can obtain repeated results by merely setting the random number generator. In addition, normalization is crucial for enhancing the model's efficacy. A normalization function was incorporated to prepare the city's weather data for training within the range 0 to 1.

For weather forecasting to anticipate the next day's data, at least one week of data is required, and it is crucial to evaluate how many timestamps (in this case, days) the LSTM model should examine before forecasting. The final phase of the initial criteria is to divide the data into training and testing sets based on the number of timestamps. 75% of validation data is used for training and 25% for testing. Depending on historical data and the required number of training epochs, the number of training days guarantees that the model learns multiple patterns. A little more than 25 epochs were adequate to train the model without overfitting, as we had previously tested this model. The testing data represents the number of days we wish to forecast. In addition, median and Gaussian filters were implemented on the dataset. The data must be converted back to its original values depending on the normalization procedure.

Data visualization is an art form that aims to display data so that non-technical audiences can better perceive and assess it. Using the Python basic module matplotlib. Pyplot, the outputs of the weather forecasting LSTM model are displayed as a series of graphs, including scatter and residual plots, training curve, predicting data (n days), and predicting data (first 75 days). In addition to these series plots, a visual and interactive dashboard was developed utilizing Power BI software to present the city's meteorological data.

Results and Discussion

Deep neural networks have successfully performed numerous machine learning tasks. By employing several processing layers and non-linear transformations, deep neural networks are able to imitate high-level abstractions in data. Recurrent neural networks (RNNs) are a type of neural network that can describe interactions across time. In contrast to conventional neural networks, RNNs employ units with internal states capable of storing information about past events. RNNs are thus suited for problems needing sequential data. Due to the problem of exploding and vanishing gradients, conventional RNNs have trouble learning long-term data associations. Long Short-Term Memory (LSTM) units use input, output, and forget gates to regulate the information in a cell state, which especially avoids this issue. Single cell, stacked, and bidirectional LSTM models are evaluated to see which is the most successful. Since LSTM models have lower error rates than other models, they can be utilized for forecasting more frequently.

After forecasting the values of the city's weather variables for five years based on the original data, we tested the results of the suggested algorithm to anticipate the city's weather and found them to be quite satisfactory. The effectiveness of LSTM is measured by its precision and MSE (Mean Square Error) value. In this study, 1500 training datasets were used to train the LSTM. Using the information from the data set, training data were generated. After being trained with the provided data, the system is put to the test by predicting the temperature, humidity, and wind speed of an upcoming day, The findings were as follows:

Table 1: The Performance comparison of the three different evaluation measures between single cell LSTM, Staked and bidirectional methods, with three intermediate variable Temperature, humidity, and wind speed, mse = mean square error.

Model Type	Temperature	Humidity	Wind speed
Type (1): Stacked Cell	mse=0.31r2=0.98	mse=6.93 r2=0.97	mse=109.15 r2=0.96
Type (2): Single Cells	mse=0.52r2=0.97	mse=14.11 r2=0.94	mse=285.05 r2=0.9
Type (3): Bidirectional Cells	mse=0.3r2=0.99	mse=31.23 r2=0.87	mse=284.97 r2=0.9

Comparison analysis of three models under study

The analysis demonstrates that the projected values and training values are very congruent, with just a little amount of error anticipated. Temperature, humidity, and wind speed are the three intermediate variables with greater than 90% forecast accuracy. The best validation accuracy was attained for temperature, with a score of 0.98 and an MSE of 0.31 for stacked model prediction. The second-best intermediate data is the humidity variable, which has a validation accuracy of 0.97 and an MSE value of 6.93. Both of these values reflect the standard error of the data. All the variables got a lower error rate and a high accuracy rate (Table 1; Figure 1 and 2).

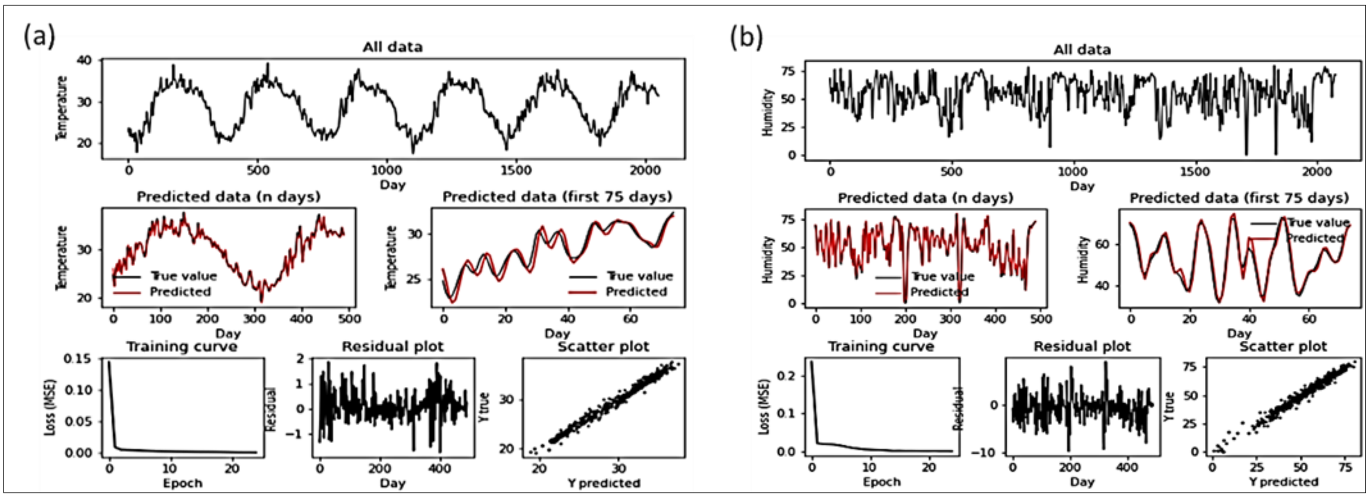

Fig 1: (a)The graphical representation of the prediction of the Temperature test data set. (b) The graphical representation of the prediction of the Humidity test data set. The black line-chart (Stacked prediction), red line-chart (true value). The smaller the error rate with a high accuracy rate, the nearer the red line is to the black line. As it shown all the lines are close to each other in all the variables.

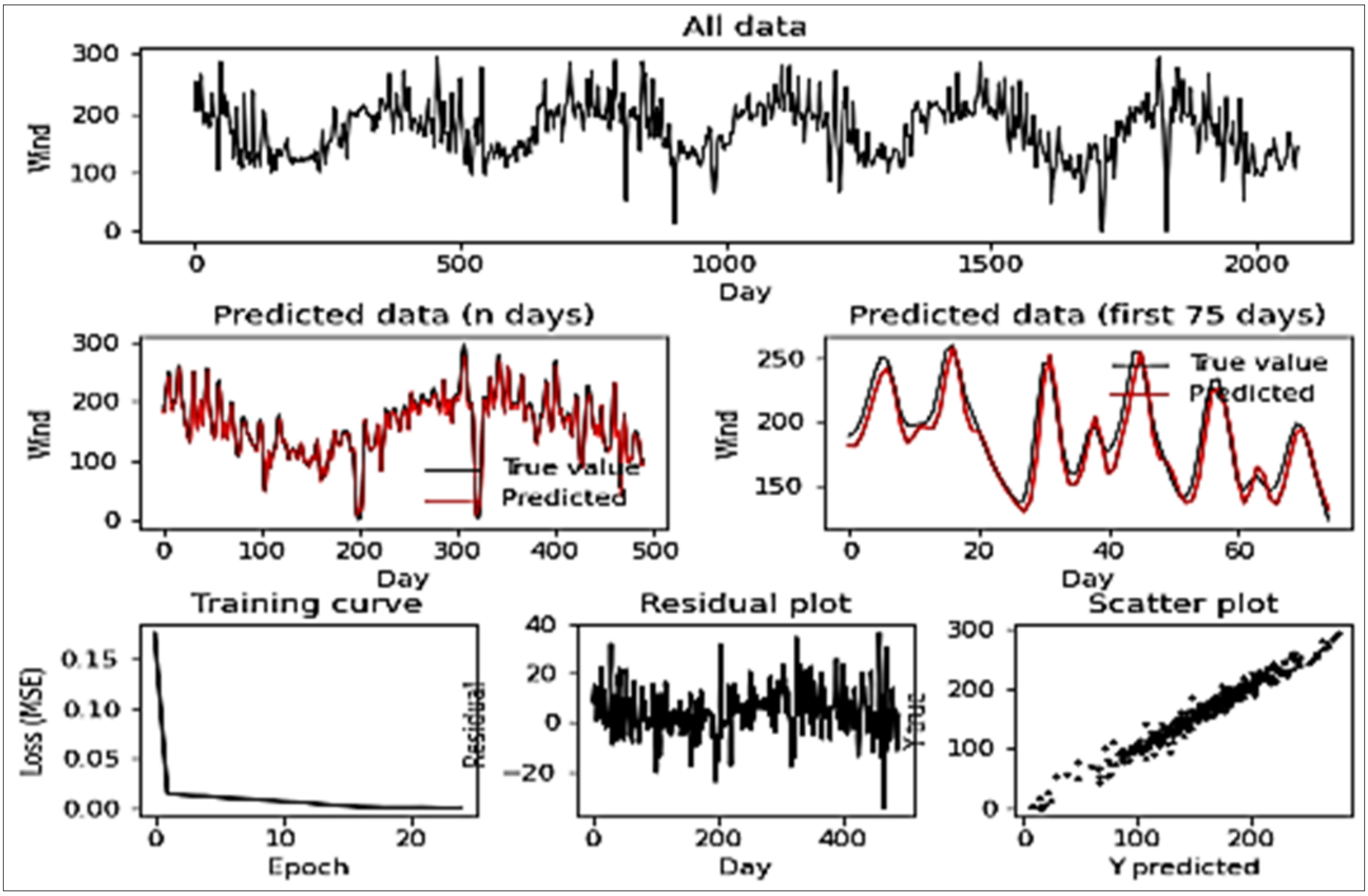

Fig 2: The graphical representation of the prediction of the Wind test data set. The black line-chart (stacked model prediction), red line-chart (true value)

The MSE value for the wind speed validation is 109.1, while the validation accuracy for the wind speed is 0.96. The predictions made by stacked LSTM were more accurate than those made by single cells and the bidirectional model (Table 1).

According to the information concerning the temperature, the humidity, and the wind speed, the Mean Squared Error (MSE) utilizing single cell models has an average value of 0.52, 14.11, and 285.05, respectively (Table 1). However, the wind speed variable had the lowest accuracy validation, with a validation accuracy of 0.90, and the humidity variable had a validation. accuracy of 0.94. When compared to the other three intermediate variables, the temperature variable had the highest accuracy validation, with a validation accuracy of 0.97 (Figure 3 and 4).

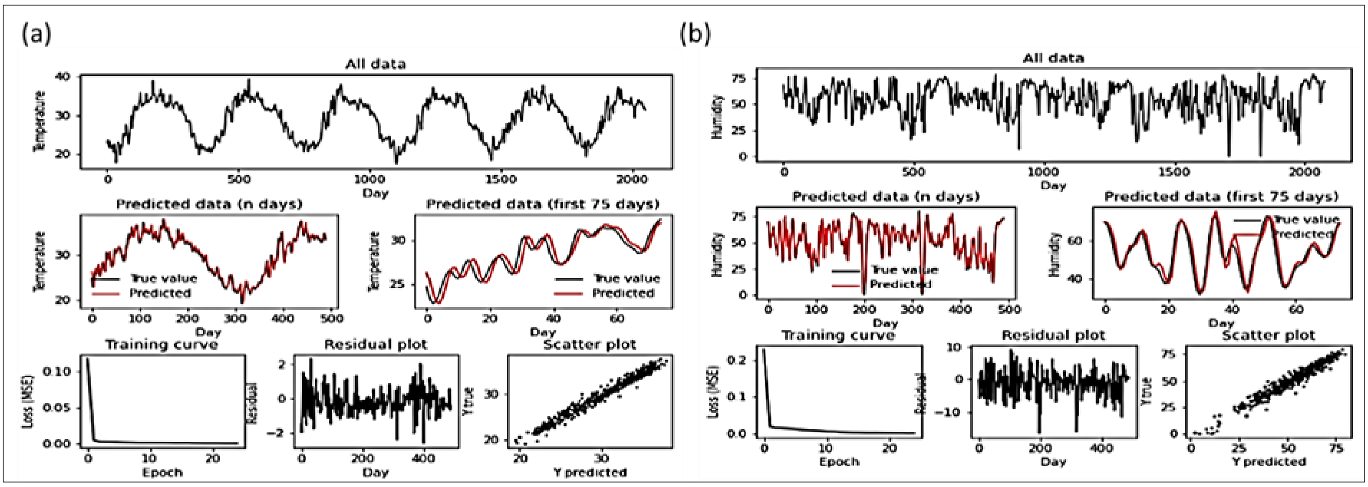

Fig 3: (a)The graphical representation of the prediction of the Temperature test data set. (b) The graphical representation of the prediction of the Humidity test data set. The black line-chart (single cell model prediction), red line-chart (true value).

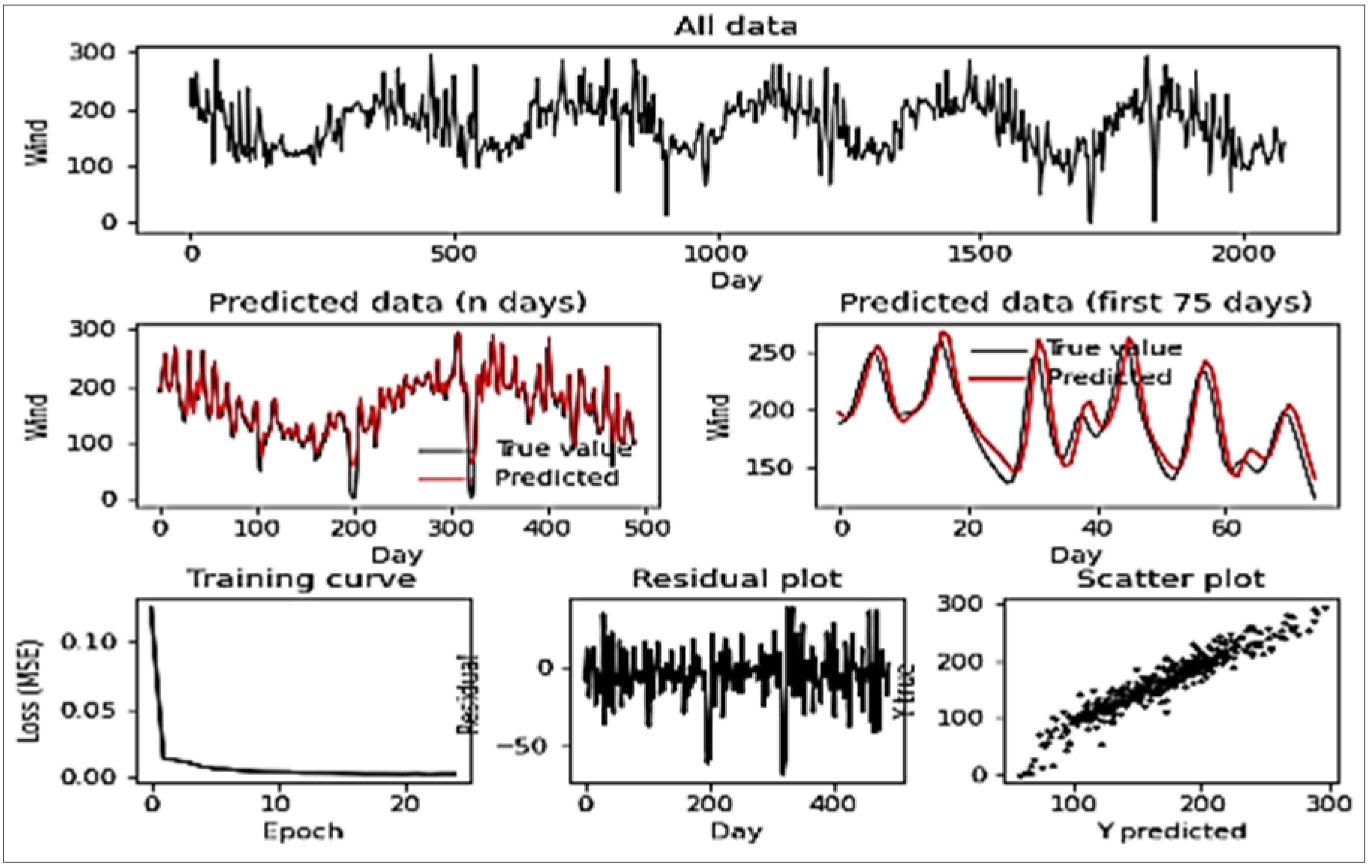

Fig 4: The graphical representation of the prediction of the Wind test data set. The black line-chart (single cell model prediction), red line-chart (true value).

The bidirectional method performed extraordinarily well in terms of temperature accuracy and wind speed accuracy, hitting 0.99 with the lowest error rate of 0.3 for temperature accuracy and 0.90 with 284.97 MSE for wind speed accuracy (Figure 5). On the other hand, it had a low accuracy rate for determining humidity, recording 0.84 with 31.23 MSE, as can be seen in figure 6 where the black line and the red line have only a minor gap between them.

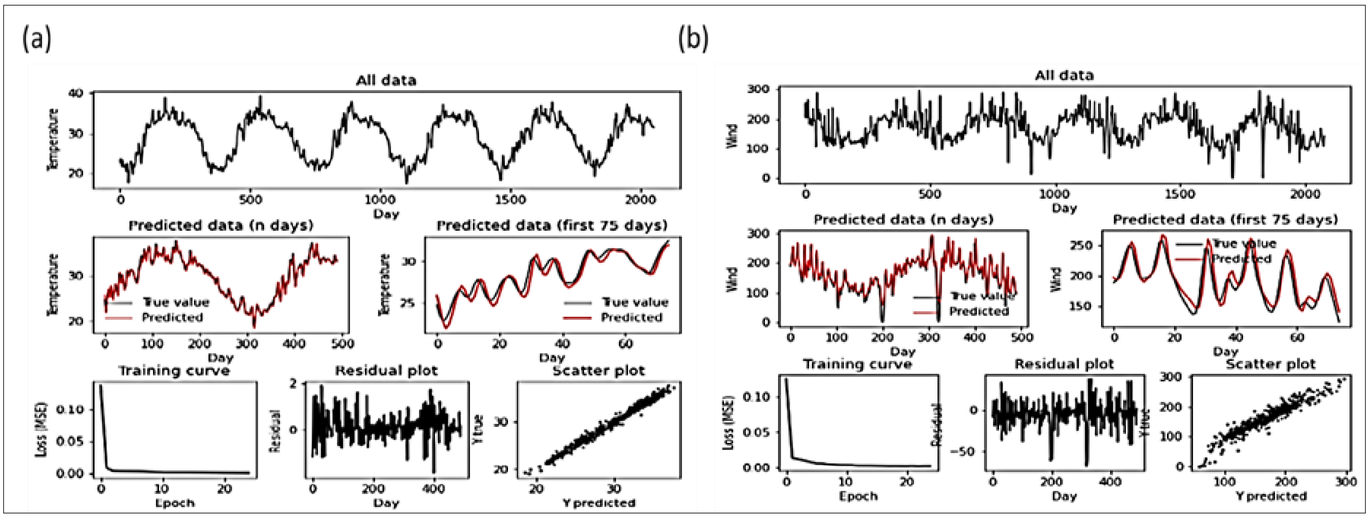

Fig 5: (a) The graphical representation of the prediction of the Temperature test data set. (b) The graphical representation of the prediction of the Wind test data set. The black line-chart (bidirectional model prediction), red line-chart (true value).

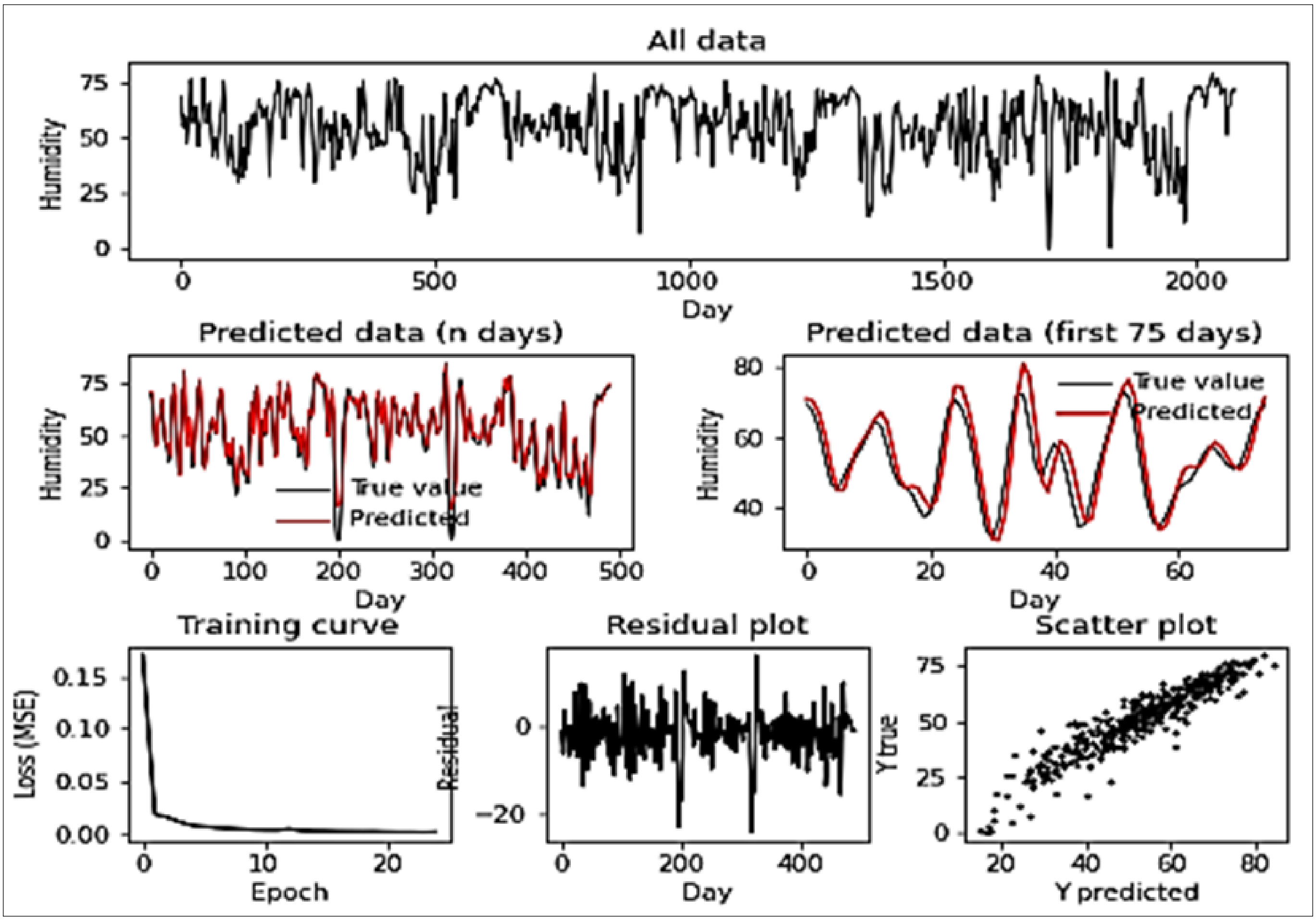

Fig 6: The graphical representation of the prediction of the Humidity test data set. The black line-chart (bidirectional model prediction), red line-chart (true value).

Conclusion

Deep neural networks are able to imitate high-level abstractions in data. This research demonstrates the efficacy of using a deep LSTM network to predict broad categories of meteorological variables. The success of the model suggests it could be applied to other weather-related problems; furthermore, python provides an excellent environment in which to compile and train models, as well as the ability to transport those models to a production server and integrate them in pre-existing applications (for instance, one could perform real-time predictions on top of an existing web application).

A new and accurate weather forecasting system that can outperform and surpass the existing conventional ones without the human intervention could be founded on a combination of numerical models and image recognition ones (in satellite photos, for instance). The predictions made by stacked LSTM were more accurate than those made by single cells and the bidirectional model.

References

1. Shrivastava G, Karmakar S, Kumar Kowar, Guhathakurta P. "Application of Artificial Neural Networks in Weather Forecasting: A Comprehensive Literature Review," *Int J Comput Appl*,2012:51:181-29. Aug., doi: 10.5120/8142-1867.
2. Akram M, El C. "Sequence to Sequence Weather Forecasting with Long Short-Term Memory Recurrent Neural Networks," *Int J Comput Appl*,2016:143:11:7-11. doi: 10.5120/ijca2016910497.
3. ewage P *et al.* "Temporal convolutional neural (TCN) network for an effective weather forecasting using time-series data from the local weather station," *Soft comput*,2020:24(21):16453-16482. doi: 10.1007/s00500-020-04954-0.
4. Abhishek K, Singh MP, Ghosh S, Anand A. "Weather Forecasting Model using Artificial Neural Network," *Procedia Technology*,2012:4:311-318. 10.1016/j.protcy.2012.05.047.
5. Ghamariadyan M, Imteaz MA, Mekanik F. "A hybrid wavelet neural network (HWNN) for forecasting rainfall using temperature and climate indices," in *IOP Conference Series: Earth and Environmental Science*, 2019, 351(1). doi: 10.1088/1755-1315/351/1/012003.
6. Chen SM, Hwang JR. "Temperature prediction using fuzzy time series," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*,2000:30:(2)263-275. doi: 10.1109/3477.836375.
7. Wong KMK, Chan PW. "Oscillatory-based Neural Network to LIDAR Data," 2008.
8. Saul Dobilas. "LSTM Recurrent Neural Networks — How to Teach a Network to Remember the Past | by Saul Dobilas | Towards Data Science," *LSTM Recurrent Neural Networks*, 2022. https://towardsdatascience.com/lstm-recurrent-neural-networks-how-to-teach-a-network-to-remember-the-past-55e54c2ff22e (accessed Oct. 07, 2022).
9. Ma M, C Liu, R Wei, Liang B, Dai J, "Predicting machine's performance record using the stacked long short-term memory (LSTM) neural networks," *J Appl Clin Med Phys*,2022:23(3):22. doi: 10.1002/acm2.13558.
10. Aad G *et al.*, "Artificial Neural Networks on FPGAs for Real-Time Energy Reconstruction of the ATLAS LAr Calorimeters," *Comput Softw Big Sci*, 2021, 5. doi: 10.1007/s41781-021-00066-y.
11. Lee C. "Understanding Bidirectional RNN in PyTorch | by Ceshine Lee | Towards Data Science,"2017. https://towardsdatascience.com/understanding-bidirectional-rnn-in-pytorch-5bd25a5dd66 (accessed Oct. 07, 2022).