International Journal of Multidisciplinary Studies and Innovative Technologies

e-ISSN: 2602-4888
dergipark.org.tr/en/pub/ijmsit
Research Article

2022, 6 (1), 113-116
DOI: DOI: 10.36287/ijmsit.6.1.113
Received: June 5, 2022; Accepted: July 17, 2022

FLOOD FORECASTING USING NEURAL NETWORK: APPLYING THE LSTM NETWORK IN THE MOSUL REGION. IRAQ

ABDULLAHI ABDU IBRAHIM[1*], AYAD KHALAF JIRRI HALBOOSH [2]

[1]Computer Engineering Department, Institute of Graduate Studies, ALTINBAS UNIVERSITY, Istanbul, Turkey
[2]Information Technology Department, Institute of Graduate Studies, ALTINBAS UNIVERSITY, Istanbul, Turkey
[*]Corresponding author: Abdullahi.ibrahim@altinbas.edu.tr

Abstract – Flooding is one of the most dangerous natural causes that inflict harm to both life and property on a yearly basis. Therefore, building a flood model for predicting the immersion zone in a watershed is critical for decision-makers. Floods are a perilous tragedy that annually threatens Iraq and the Middle East region, impacting millions of people. In this context, having suitable flood forecasting algorithms may help people by reducing property damage and saving lives by warning communities of potentially severe flooding events ahead of time. Data mining techniques such as artificial neural network (ANN) approaches have recently been applied to model floods. The purpose of this study is to develop a model that extrapolates the past into the future using existing statistical models and recurrent neural networks and is powered by rainfall forecasting data. We investigate a number of time series forecasting approaches, including Long Short-Term Memory (LSTM) Networks. The forecasting methods investigated are tested and implemented using rainfall data from the Mosul region of Iraq. In addition, in flood occurrences and conducting experiments to study the relationship between rainfall and floods.

Keywords *– Artificial-Neural-Network, Long Short-Term Memory, Rainfall, Forecasting, Floods*

Citation: Ibrahim, A.A., Halboosh, A.K.J. (2022). Flood Forecasting Using Neural Network: Applying the LSTM Network in the Mosul Region. Iraq. International Journal of Multidisciplinary Studies and Innovative Technologies, 6(1): 113-116.

I. INTRODUCTION

The study's background is that the geography of Mosul, Iraq, features specific patterns of flood vulnerability that have contributed to intensifying and different patterns of flooding. This study is applied to one place, Mosul, as a proof of concept, although the results might be more beneficial for flood predictions in general.

A flood estimate exhibit will be critical in displaying essential facts about the possibility of unavoidable flooding in populous areas. By constructing such models, disruption in ranges like Mosul may be reduced by reducing the financial and environmental costs of floods. More precisely, the prediction approach designed for South Asia, particularly the Mosul region, will reduce the likelihood of suffering and tragedy in life. If applying LSTM's artificial neural network (ANN) models can provide realistic nitty-gritty expectations, the lead time for flood caution may be increased by one day, and the subsequent flood emergency procedures can be encouraged and boosted.

The primary objective of this research is to investigate and develop bogus neural systems using LSTM Arrange that can be used as a demonstration in an environment such as the city of Mosul as shown in figure (1) to estimate the commencement of floods. Several types of constructed neural network models are examined in order to determine the neural organization characteristics that will have the best estimate for unavoidable floods, including their designs and modifications of associated learning laws.

Fig. 1 Map of Iraq

number of sub-goals were established in order to reach this conclusion. The sub-goal is to show how artificial neural networks (LSTM Arrange) can be used to estimate floods using meteorological data in a reliable and lucrative manner. To define the neural organization architecture that will deliver the greatest anticipated precipitation consequences.

II. MATERIALS AND METHOD

The proposed LSTM flood assessment (LSTM-FF) system is composed of multivariate single-step LSTM systems that acknowledge spatial and common component data from real and anticipated precipitation and early release as inputs. The Mosul case considered, it appears that the proposed models can expect streak surges, especially enormous surge events, with high precision (the degree of the number of qualified occasions to the common number of surge occasions).

Small floods can assist the LSTM-FF consummate in investigating a better rainfall-runoff link for large flood simulations. According to the load impact study in LSTM network topologies, the discharge input has a bigger influence on the 1-hour LSTM network and this influence diminishes with lead-time. Meanwhile, the LSTM networks investigated a similar association in the prior lead-time.

The LSTM surge determining (LSTM-FF) show, which is made up of T multivariate single-step LSTM systems, is outlined for figure release with a 1 hr. lead-time. Each release is anticipated employing a particular LSTM arrangement, as seen in Figure (1). Observations: a. show and past (1 hour, 2-hour, H hour lag) precipitation at each rain station; b. prior discharge at the outlet; and (2) gauge short-term expected precipitation with a T-hour lead-time.

The observed inputs are represented by X_{t-H}, ..., X_{t-1}, X_t, while the forecasting inputs are represented by X, X, ..., X. The outputs include t+1 t+2 t+T forecast discharge $q^{sim, qsim}$ ...,q^{sim} rectified Linear Unit (ReLU) with values in the range (0, +∞) is used to output non-negative values.

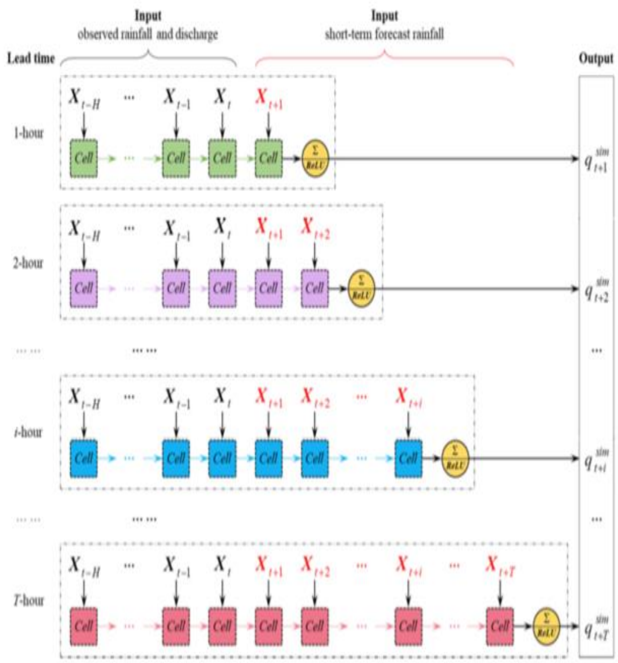

Fig. 2 Structure of the LSTM flood forecasting model

The following are the steps for model training.

Stage 1: Define the release lead-time (T) based on the realistic need for streak spike early warning within the designated watershed. **Stage 2:** Create and standardize a data set. Because precipitation and release have obvious physical consequences and quantities, the data is balanced by using these conditions.

$$\hat{x} = \frac{x - x_{min}}{x_{max} - x_{min}}$$

where $x^{\exists}$ is the normalized value, x denotes the observed value, x_{max} and x_{min} denote the maximum and minimum observed values.

Stage 3: Divide the information set into three sections: preparation, approval, and testing. **Stage 4:** In that sequence, set the hyperparameters (units, bunch estimate, and age) and prepare the 1-hour, 2-hour, and T-hour LSTM systems. C t and h t measurements are presented in units. Batch size determines the number of tests that will be broadcast through the arrangement. The number of passes the machine learning computation has performed across the entire prepared dataset is represented by an age. When the clump estimate is the whole prepared dataset, the clump measure and age are the same. To determine the initialization of weights, an are **Stage 5:** Repeat step 4 using trial and error to compute the ultimate esteem of hyperparameters for 1-hour, 2-hour, and T-hour LSTM systems independently. The learning curve is used to keep a strategic gap between being over- or under-fitting.

Stage 6: Repeat stage 4 using trial and error to compute the ultimate esteem of hyperparameters independently for 1-hour, 2-hour, and T-hour LSTM systems. The learning curve is used to keep a strategic gap between being over-or under-fitting.

Stage 7: Save the best demonstration based on your trial-and-error discoveries from step 6. Apply the previously saved LSTM-FF demonstration to the input test set and normalize the yield to the reenacted releases. **Stage 8:** Examine the LSTM-FF model's reenacted results.

For exploratory preparation reenactments, the day-by-day crude precipitation estimations from an assortment of meteorological stations in and around Fort Worth were utilized. Thirteen years of data (from 2007 to 2109) on daily precipitation records obtained from www.meteoblue.com, where we conducted our research. As the beginning set of exploratory information, the area known as Al-Faisaliyah was chosen. This area was chosen since it's in a locale that's generally close to the Tigris Waterway.

The panda's data frame is a structure that stores two-dimensional data and the labels that go with it. Data-Frames are widely used in data science, machine learning, scientific computing, and a variety of other data-intensive disciplines. Data-Frames are comparable to SQL tables or spreadsheets used in Excel or Calc. Because they are an essential part of the

Python and NumPy ecosystems, Data-Frames are often quicker, easier to use, and more powerful than tables or spreadsheets. The dataframe.sum () method in Pandas returns the sum of the values for the specified axis. If the input is an index axis, it adds all the values in a column and then continues the process for all columns, returning a series with the total of all the values in each column. It also allows you to bypass missing values in the data-frame when computing the total. The data type of the date column was one of the key challenges that arose. The approach was to convert it to a timestamp and then back to Date-Time format.

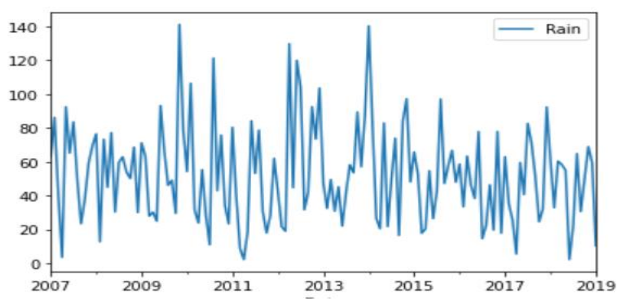

Fig. 3 presenting the data in a plot figure

III. RESULTS

Observe that raising the model's layers marginally improves performance; LSTM layers following the top layer are started with the hidden states and cell states of the preceding LSTM layer. In general, the internal states of each LSTM layer are randomly initialized.

The following period in deciding the execution of the LSTM demonstrate is to match precipitation information from a few long time periods as well as the expected and measured values within the test dataset. Precise surge estimating will serve as a foundation for flood risk administration, planning, and management.

The expected surge crests rise at the same time as the real maxima in this situation. On the occasion of estimating the most extreme flowrate one day and two days ahead, a satisfactory resilience esteem of almost 5% and 14%, respectively, is utilized within the hydrological forecast. This is often due to the trouble of precisely anticipating the greatest discharge value during the surge season, especially in places with complicated landscapes and soaking slopes like the ponder zone. Moreover, the figures (3) suggest that the LSTM could be a data-driven show established on factual associations among input and yield information in all forecast occurrences. As a result, one of the basic criteria impacting the model's exactness is the relationship between the information arrangement at the target-forecast location and further situations.

IV. DISCUSSION

The inquire about given over illustrates the LSTM model's fabulous advantage in its capacity to effectively learn brief conditions, and this demonstration can totally be It is used to forecast the flow two or three days ahead of time with an accuracy of more than 86 percent. Of their effortlessness, data-based procedures are a practical approach with high precision to modeling the hydrological readiness, particularly in developing countries such as Vietnam, where the application of blocked off identifying development inside the improvement of a real-time surge caution system is compelled.

The figure (4) shows that the actual percent rainfall increased, while our model projected that it would increase as well. This clearly demonstrates how effective LSTMs are in analyzing time series and sequential data.

As we can see in figure (4) As we can see, the blue continuous line represents the real data of the rainfall in all the periods of time across the area that it has been represented. It manifests itself in various ways and at various times. Then comes the orange line, which shows the predation of the flood. We are all aware that more rainfall means more flooding. The case of the flood is higher in this graph, where it is more than 2.0. And the graph demonstrates that it is continuous; every time the amount of rainfall goes up in a small area, the orange line appears, showing that the possibility of a flood may be realistic or it will happen. The graphic shows that the actual percent of rainfall increased, while our model projected that it would increase as well. The expected flood line rises at the same time as the real maxima in this situation. On the occasion of estimating the most extreme flowrate one day and two days ahead, a satisfactory resilience index of almost 5% and 14% was achieved, respectively.

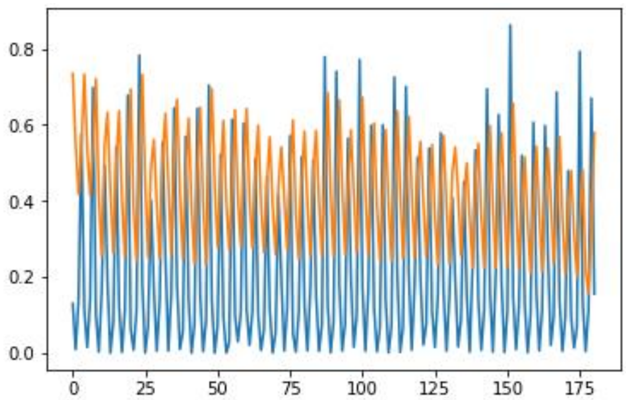

Fig. 4 The Result of LSTM

V. CONCLUSION

Based on the data-driven procedure, this is how they advertised a successful arrangement for surge estimating. To illustrate, the LSTM neural organization show was created and extensively tested. The LSTM show has learnt long-term associations between successive information arrangements and has been demonstrated to be solid in surge estimating.

In spite of the fact that the LSTM show effectively addresses successive information issues, there are different confinements that must be considered. LSTM models, in common, are data-driven models that, like physical-based models, result in destitute modeling of hydrological forms. LSTM (or ANN)-based models in particular, because they provide extremely precise estimates at specified locations within the investigated area. As a result, these models ought to be coordinated with meteorological models such as precipitation estimating models to make strides in long-term expectation execution.

ACKNOWLEDGMENT

For the ancestors who paved the path for us upon whose shoulders we stand. This is also dedicated to my family and friends who supported me on this journey. Thank you I would like to express my deepest gratitude to my supervisor Dr. ABDULLAHI ABDU IBRAHIM for his unwavering support,collegiality and mentorship throughout this research.

.

REFERENCES

[1] S. Mei, A. Montanari, and P.-M. Nguyen, "A mean field view of the landscape of two-layer neural networks," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 115, no. 33, pp. E7665–E7671, 2018.

[2] V. Gulshan *et al.*, "Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs," *JAMA*, vol. 316, no. 22, p. 2402, 2016.

[3] Y. Liu, S. Liu, Y. Wang, F. Lombardi, and J. Han, "A survey of stochastic computing neural networks for machine learning applications," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 7, pp. 2809–2824, 2021.

[4] J. Bayer, D. Wierstra, J. Togelius, and J. Schmidhuber, "Evolving memory cell structures for sequence learning," in *Artificial Neural Networks – ICANN 2009*, Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 755–764.

[5] G. Bellec, D. Salaj, A. Subramoney, R. Legenstein, and W. Maass, : "Long short- term memory and learning-to-learn in networks of spiking neurons," 2018, pp. 787–797.

[6] H. Gao, J. Mao, J. Zhou, Z. Huang, L. Wang, and W. Xu, : "Are you talking to a machine? dataset and methods for multilingual image question," 2015, pp. 2296–2304.

[7] J. Gong, X. Chen, T. Gui, and X. Qiu, : "Switch-lstms for multi-criteria chinese word segmentation," 2019, vol. 33, pp. 6457–6464.

[8] K. Greff, R. K. Srivastava, J. Koutnik, B. R. Steunebrink, and J. Schmidhuber, "LSTM: A search space odyssey," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 10, pp. 2222–2232, 2017.

[9] C. R., K. K. R., and D. E. Newton, "Drthis: Deep ransomware threat hunting and intelligence system at the fog layer," *Future Generation Computer Sys- tems*, vol. 90, pp. 94 – 104, 2019.

[10] T. Horsmann and T. Zesch: "Do lstms really work so well for pos tagging? – a replication study," 2017, pp. 727–736.

[11] Cho, K.; van Merrienboer, B.; Gülçehre, Ç.; Bougares, F.; Schwenk, H.; Bengio, Y. Learning phrase representations using RNN encoder-decoder for statistical machine translation. arXiv 2014, arXiv:1406.1078.