Time series forecasting for uni-variant data using hybrid GA-OLSTM model and performance evaluations

Nisha Thakur · Sanjeev Karmakar · Sunita Soni

Abstract Time series forecasting of uni-variant rainfall data is done using a hybrid genetic algorithm integrated with optimized long-short term memory (GA-OLSTM) model. The parameters included for the valuation of the efficiency of the considered model, were mean square error (MSE), root mean square error (RMSE), cosine similarity (CS) and correlation coefficient (r). With various epochs like 5, 10, 15 and 20, the optimal window size and the number of units were observed using the GA search algorithm which was found to be (49, 9), (12, 8), (40, 8), and (36, 2) respectively. The computed MSE, RMSE, CS and r for 10 epochs were found to be 0.006, 0.078, 0.910 and 0.858 respectively for the LSTM model, whereas the same parameters were computed using the Hybrid GA-OLSTM model was 0.004, 0.063, 0.947 and 0.917 respectively. The experimental results expressed that the Hybrid GA-OLSTM model gave significantly better results comparing the LSTM model for 10 epochs has been discussed in this research article.

Keywords Forecasting · Long short-term memory (LSTM) · Recurrent neural networks (RNN) · Artificial neural network (ANN) · Deep learning · Uni-variant time series · Genetic algorithm (GA)

1 Introduction

Forecasting rainfall using a uni-variant dataset is almost impossible due to its internal variability and chaotic nature, however, some approaches have provided prominent results [1, 2]. Nowadays various approaches have been used for prediction, based on statistical and soft computing techniques [3]. Accordingly, various techniques based on machine learning, including artificial neural network (ANN) [4–7], support vector machine (SVM) [8], adaptive basis function (ABF) [9], multiple layer perceptron(MLP) [10], multiple nonlinear regression (MNLR) [11], adaptive neuro-fuzzy inference system (ANFIS) [12] have been used extensively. Deep learning, which has the superior ability to reflect nonlinearity and complex characteristics of rainfall time series, is also being used nowadays for the prediction of rainfall. The dynamical and the empirical approach are the fundamental approaches in deep learning. In the dynamical approaches, physically built models were used for predictions that are entirely dependent on the system’s equation that predicts the rainfall, whereas the other method i.e. the empirical scheme is dependent on the analysis of historical weather data and their correspondence with various atmospheric variables over multifarious regions. Regression, fuzzy logic and deep learning techniques are the most widely employed within the empirical scheme [13–21].

However, with LSTM, poorly selected initial weights could sometimes lead to inaccurate performance. The inspiration behind this study was to consider a method such that LSTM could start with a set of properly selected initial weightings to reduce forecasting error. Specifically, the considered hybrid approach (GA-OLSTM) is responsible for both the learning and memorization intelligence using the LSTM Network along with the searching ability of the
Salman et al., proposed a merged-LSTM model and a comparison, was made with the LSTM model. The author reported the merged LSTM model worked well in predicting visibility by 4.8% higher accuracy [14]. Poornima and Pushpalatha informed an LSTM (intensified) based recurrent neural network (RNN) to predict rainfall. The results obtained by them were compared with various models to evaluate the improvement to forecast rainfall [23]. Crivellari and Beinat proposed another deep learning approach based on a problem related to trajectory prediction. The author reported that the LSTM model outperformed in comparison with global Markov model (GMM) and variable-order Markov model (VGMM) [24]. Ding and Qin proposed a multi-value associated network model of LSTM-based on a deep recurrent neural network (Associated Net). Experimental results proved that the average accuracy of the Associated Net model was better than that of the other two models discussed and it has the ability to predict the multiple values together. They reported an accuracy of more than 95% in each of the predictions [25]. Qiu et al. compared the results obtained using their proposed model with three other models including the LSTM model, the LSTM with wavelet denoising and the gated recurrent unit (GRU) neural network model for prediction of stock and they also reported that LSTM with attention model performs outstandingly in comparison with other models [26]. Chimmula and Zhang presented LSTM net-model performs outstandingly in comparison with other models [26]. Ouma et al. compared the result between LSTM and wavelet neural network for rainfall prediction and runoff trend analysis [28].

2 Dataset and algorithm used

In this work, we required to collect time series monthly rainfall historical data of Chhattisgarh region in order to predict the rainfall in coming months. Forecasting of monthly rainfall was planned using the hybrid approach (GA-OLSTM). This algorithm was implemented using Keras and Tensor flow libraries of Python 3.9.7 software.

2.1 Collection of time series monthly dataset

A time-series uni-variant monthly rainfall data of Chhattisgarh State, India from January 1901 to December 2017 was collected. In all, monthly rainfall data of 1404 months were used and has been collected from the Indian Meteorological Department (IMD). The geographical location of Chhattisgarh, India is 17°46′ north to 24°5′ north latitude and from 80°15′ east to 54°20′ east longitude [29].

2.2 LSTM model

LSTM is considered a technique to remove the problem of long-term dependency. LSTM networks can have historical information of more than thousand-time steps. LSTM is possible to be scaled to a lot longer sequences than a traditional RNN, removing the inbuilt drawbacks of a traditional RNN, i.e., exploding and vanishing gradients. Nowadays, LSTM is highly used in many sequential modelling operations, like speech recognition, natural language processing and motion detection [19, 30]. The LSTM is a deep recurrent neural network-based model consisting of a memory cell and incorporating, three multiplicative gating units input gate, output gate and a forget gate. A recurrent connection is being established between the cells. Every gate does continuous operations for the cells like write, read and reset. The cell is for conveying “state” values over arbitrary time intervals [30–32]. The calculation procedure for the Block of LSTM block is been discussed earlier [19, 30–33, 36, 37].

2.3 Hybrid GA-OLSTM model

The considered work consists of an LSTM network integrated with a GA search algorithm. As previously elaborated, the initial values of the weighting matrices may make the difference in the results obtained by LSTM [19]. GA has been used to assist in searching the appropriate initial values for the weighting matrices of the LSTM. GA is a searching algorithm based on population that utilizes a chromosomes population in the process of searching. Every chromosome speaks for a suitable solution. Figure 1 exhibits the flow diagram of the method used.

In this study, a hybrid approach of LSTM network combined with GA for finding the appropriate time window and number of LSTM units for rainfall time series forecasting. Since the past information is used in the learning process in the LSTM network, an appropriate window size becomes an important factor to be considered. It is to note that, for the window a very small window size, important information will be neglected by the model whereas, for very large window size, the model will be over-fitted on the training data.

The evolutionary search algorithm, GA was used, to identify the optimal size of time windows and the number of LSTM units. The GA can be split into various stages such as initialization stage, fitness calculation stage; termination criteria check stage, selection stage, crossover stage, mutation stage, evaluation and replacement etc. [21, 33, 35]. In this study, binary arrays are used as
chromosomes and the MSE of the prediction model was functioning like the fitness value.

In the second stage of the algorithm, to evaluate the fitness of GA various sizes of time windows along with different numbers of LSTM units are applied. The populations are initialized with random values that consist of possible solutions, before the start of exploring the search space by the genetic operators. The chromosomes taken in this study were converted into binary digits that specify the size of the time window and the number of LSTM units. The selection and recombination operators start to search for a better solution with the help of the genetic operators. A predefined fitness function is used to evaluate the solutions, and strings with excellent results opted for the reproduction. An important part of GA is the fitness function and is required to be selected with attention. In this study, RMSE was chosen to calculate the fitness of each chromosome, and various architectural factors that return the shortest RMSE was selected as the optimal solution. If the result of the reproduction process fulfils the termination criteria, derived optimal or near-optimal solution is applied to the forecasting model. The whole process is repeated, if the result does not satisfactory. To find an excellent solution for the problem, genetic parameters, such as crossover rate, mutation rate, and population size, can affect the parameter. In this work, population size, number of generations, probability of mutating an individual and probability of mating two individuals were taken as 4, 4, 0.1 and 0.4 respectively.

3 Results and discussion

A time series uni-variant dataset was used to predict monthly rainfall using LSTM model. Then GA-OLSTM model which is a hybrid model was also used in efforts to improve the prediction accuracy. The comparative analysis was done to conclude the best combination of windows size and number of LSTM units.

3.1 Experimental outcomes

In the experimental results, we started experimenting with the time series Uni-Variant data by applying the LSTM model. With 10 epochs various factors such as MSE, RMSE, CS, and r were calculated to investigate the accuracy of forecasting data and found as 0.006, 0.078, 0.910 and 0.858 respectively which is illustrated in Table 1. A deep learning tool was utilized to find these parameters. These values were an indicator of poor resemblance between observed (actual) and forecast data; therefore, the Hybrid GA-OLSTM model was taken into consideration for further experimental work. The graph obtained for window size and number of units (20, 16) is shown in Fig. 2a.

A prediction is recommended to use a hybrid (GA-OLSTM) approach of LSTM network integrated with GA to find the optimized window size and the number of LSTM units for rainfall forecasting using uni-variant time series data. The whole study was carried out in two sections. The first section of the analysis involved designing the appropriate LSTM network. An LSTM network was used with a sequential input layer followed by two hidden layers, incorporating the optimal parameters investigated by GA for 5, 10, 15 and 20 epochs respectively. As a result of GA search for 5, 10, 15 and 20 epochs, the optimal window size and number of units were observed to be (49,9), (12,8), (40,8), and (36,2) respectively.

| Table 1 LSTM network-based analysis of uni-variant time-series data |
|---------------------------------------------------------------|
| LSTM input parameter | Window size = 20, Number of units = 16 |
| Accuracy parameters | 10 epochs |
| MSE | 0.006 |
| RMSE | 0.078 |
| CS | 0.910 |
| r | 0.858 |
Subsequently, in the second section forecasting of time series uni-variant data was done using the LSTM model with the optimal parameters obtained from the results of the first section.

To examine the efficiency of the results obtained using the Hybrid GA-OLSTM model, the above used statistical-based various parameters were estimated. We went for the various window size and the number of units attained from GA search to observe the quality of prediction for various epoch(s) such as 5, 10, 15 and 20 respectively (Table 2). The factors MSE and RMSE are expected to be less for obtaining optimum prediction quality. On the other hand, the values of CS and $r$ must be close to one. When these values are closed to 1, we consider the actual and predicted data are very similar and when the said values are closed to 1, we may conclude about the usefulness of the prediction made. If we observe Table 2, it was found that the value of RMSE and MSE is 0.063 and 0.004 for 10 epochs, which is the smallest in comparison to that of prediction result obtained for the rest of the epochs. We also got the value of CS and $r$ as 0.947, and 0.917 respectively which is again highest for 10 epochs. The graph obtained for window size and number of units (12, 8) which gave the best possible result is shown in Fig. 2b. In this figure, a great resemblance between the actual and forecast data can be observed.

| LSTM input parameter | Window size = 49, number of units = 9 | Window size = 12, number of units = 8 | Window size = 40, number of units = 8 | Window size = 36, number of units = 2 |
|----------------------|--------------------------------------|--------------------------------------|--------------------------------------|--------------------------------------|
| Accuracy parameters  | 5 epochs                             | 10 epochs                            | 15 epochs                            | 20 epochs                            |
| MSE                  | 0.006                                | 0.004                                | 0.004                                | 0.005                                |
| RMSE                 | 0.079                                | 0.063                                | 0.063                                | 0.075                                |
| CS                   | 0.89                                 | 0.947                                | 0.936                                | 0.905                                |
| $r$                  | 0.821                                | 0.917                                | 0.899                                | 0.849                                |

Fig. 2 a Comparison between actual and forecast rainfall (10 epoch(s), optimal window size = 20, number of LSTM units = 16) using LSTM algorithm. b Comparison between actual and forecast rainfall (10 epoch(s), optimal window size = 12, number of LSTM units = 8) using hybrid (GA-OLSTM) algorithm.
3.2 Comparison of performance between LSTM and hybrid GA-OLSTM model

A comparison is being done with the performance of the Hybrid GA-OLSTM with the LSTM Model. As depicted in Table 3 LSTM was less stable and has large errors. On the other hand, Hybrid GA-OLSTM presented better performance. The predicted MSE and RMSE of the LSTM model is 0.006 and 0.078, while the predicted MSE and RMSE of the GA-OLSTM model is 0.004 and 0.063 respectively. Lastly, the predicted CS and r of the LSTM model is 0.91 and 0.858, while the same for the GA-OLSTM model is 0.947 and 0.917 respectively, which express great accuracy resemblance between the actual and predicted data.

| Model          | Optimal parameters | MSE  | RMSE  | CS    | R    |
|----------------|--------------------|------|-------|-------|------|
| LSTM           | Window size = 20, number of units = 16 | 0.006 | 0.078 | 0.91  | 0.858|
| Hybrid GA-OLSTM | Window size = 12, number of units = 8  | 0.004 | 0.063 | 0.947 | 0.917|

5 Future scope

The work carried out in this article may be extended by using the reported GA-OLSTM model for the multivariate data. In this work, univariate time series data was used for analysis whereas, in the future, multivariate data may be processed for obtaining better prediction results. It is planned to access the desired multivariate data from Google Earth Interface “CRU TS Version 4.04”. The interface discussed is available on the website of the Climate Research Unit (CRU).

Author contributions NT rest of article. SK supervision. SS supervision.

Funding No funding support was received.

Availability of data and material Data is collected from the website of the Indian Metrological Department Pune (https://www.springernature.com/gp/authors/research-data-policy/data-availability-statements/12330880).

Code availability Coded by ourselves in python.

Declarations

Conflict of interest There is no conflict of interest.

Ethics approval The work submitted is the original work of our team and the contents submitted above are not submitted elsewhere for publication.

Consent to participate Both the authors have contributed to the article.

Consent for publication We give our consent for the publication of identifiable details, which can include photograph(s) or details within the text (“Material”) to be published in the above Journal, therefore, anyone can read material published in the Journal.

References

1. Lin X, Yang Z, Song Y (2009) Short-term stock price prediction based on echo state networks. Expert Syst Appl 36:7313–7317
2. Adebiyi AA, Adewumi AO, Ayo CK (2014) Comparison of ARIMA and artificial neural networks models for stock price prediction. J Appl Math 2:1–7
3. Nayak DR, Mahapatra A, Mishra P (2013) A survey on rainfall prediction using artificial neural network. Int J Comput Appl 72:32–40
4. Vamsidhar E, Varma KV, Rao PS, Satapati R (2010) Prediction of rainfall using back-propagation neural network model. Int J Comput Eng 2:1119–1211
5. Abdulkadir TS, Salami AW, Kareem AG (2012) Artificial neural network of rainfall in Ilorin, Kwara State, Nigeria. USEP Res Inf Civ Eng 9(1):108–120
6. Abdulkadir T, Salami S, Aremu AW, Ayanshola AS, Oyejobi AM (2017) Assessment of neural networks performance in modeling rainfall amounts. J Res For Wildl Environ 9(1):12–22
7. Chakraborty K, Mehrotra K, Mohan CK, Ranka S (1992) Forecasting the behaviour of multivariate series using neural networks. Neural Netw 5:961–970
8. Garcia EGO, Sanz SS, Mateo CC (2014) Accurate precipitation prediction with support vector classifiers: a study including novel predictive variables and observational data. Atmos Res 139:128–136
9. Philip NS, Joseph KB (2003) A neural network tool for analysing trends in rainfall. Comput Geosci 29:215–223
10. Hashim FR, Daud NGN, Ahmed KA, Adnan J, Rizman ZI (2017) Prediction of rainfall based on weather parameter using artificial neural network. J Fundam Appl Sci 9:493–502
11. Rajashree A, Pal J, Das D, Chaudhuri S (2018) Long range forecast of Indian summer monsoon rainfall using artificial neural network model. Meteorol Appl 26:347–361
12. Solgi A, Nourani V, Pouraghaj A (2014) Forecasting daily precipitation using hybrid model of wavelet-artificial neural network and comparison with adaptive neuro-fuzzy inference system (case study: Verayneh Station, Nahavand). Adv Civ Eng 1–12
13. López E, Carlos Valle C, Allende H, Gil E, Madsen H (2018) Wind power forecasting based on echo state networks and long short-term memory. Energies 11:526.
14. Salman AG, Heryadi Y, Abdurahman E, Suparta W (2018) Weather forecasting using merged long short-term memory model. Bull Electr Eng Inform 7(3):377–385
15. Chung H, Shin K (2018) Genetic algorithm-optimized long short-term memory network for stock market prediction. Sustainability 10:3765. https://doi.org/10.3390/su10103765
16. https://en.wikipedia.org/wiki/Rice_production_in_India. Accessed 25 Sept 2020
17. Chakraborty S, Pandey RP, Chaube UC, Mishra SK (2013) Trend and variability analysis of rainfall series at Sonath River Basin, Chhattisgarh (India). Int J Appl Sci Eng Res 2(4):425–434
18. Bhuarya SK, Chaudhury JL, Khalkho M, Khalkho D (2015) Comparison of drought indices at different stations of Chhattisgarh. J Agric Phys 15(2):140–149
19. Thakur N, Karmakar S (2021) Deep learning approach using long short term memory technique for monthly rainfall prediction in Chhattisgarh, India. Int J Sci Res Comput Sci Eng 9:8–13
20. Adjei OC et al (2021) Rainfall forecasting in sub-Saharan Africa-Ghana using LSTM deep learning approach. Int J Eng Res Technol 10(3)
21. Kim HJ, Shin KS (2007) A hybrid approach based on neural networks and genetic algorithms for detecting temporal patterns in stock markets. Appl Soft Comput 7:569–576
22. Santra SA, Lin LJ (2019) Integrating long short-term memory and genetic algorithm for short-term load forecasting. Energies 12:2040. https://doi.org/10.3390/energies12112040
23. Poornima S, Pushpalatha M (2019) Prediction of rainfall using intensified LSTM based recurrent neural network with weighted linear units. Atmosphere 10:668. https://doi.org/10.3390/atmos10110668
24. Crivelli A, Beinat E (2020) LSTM-based deep learning model for predicting individual mobility traces of short-term foreign tourists. Sustainability 12:349. https://doi.org/10.3390/su120310349
25. Ding G, Qin L (2020) Study on the prediction of stock price based on the associated network model of LSTM. Int J Mach Learn Cybern 11:1307–1317
26. Qiu J et al (2020) Forecasting stock prices with long-short term memory neural network based on attention mechanism. PLoS ONE 15(1):e0227222
27. Chimmula VKR, Zhang L (2020) Time series forecasting of COVID 19 transmission in Canada using LSTM networks. Chaos Solitons Fractals 135:109864
28. Ouma YO et al (2022) Rainfall and runoff time-series trend analysis using LSTM recurrent neural network and wavelet neural network with satellite-based meteorological data: a case study of Nzoia hydrologic basin. Complex Intell Syst 8(1):213–236
29. https://data.gov.in/sites/default/files/datafile/sub-division_rainfall_act_dep_1901-2015.csv. Accessed 30 Aug 2021
30. Good Fellow I, Bengio Y, Courville A, Bengio Y (2016) Deep learning. MIT Press, Cambridge, pp 373–418
31. Schmidhuber J, Hochreiter S (1997) Long short-term memory. Neural Comput 9:1735–1780
32. Gers FA, Schmidhuber J, Cummins F (1999) Learning to forget: continual prediction with LSTM. Neural Comput 12:2451–2471
33. Kim Y, Roh JH, Kim H (2017) Early forecasting of rice blast disease using long short-term memory recurrent neural networks. Sustainability 10:34–53
34. Armano G, Marchesi M, Murr M (2005) A hybrid genetic-neural architecture for stock indexes forecasting. Inf Sci 170:3–33
35. Pal SK, Wang PP (1996) Genetic algorithms for pattern recognition. CRC Press, Boca Raton, p 336
36. Divate MS (2021) Sentiment analysis of Marathi news using LSTM. Int J Inf Technol 13:2069–2074
37. Kumar R, Kumar P, Kumar Y (2021) Multi-step time series analysis and forecasting strategy using ARIMA and evolutionary algorithms. Int J Inf Technol. https://doi.org/10.1007/s41870-021-00741-8