Applications of Deep Learning Models for Forecasting and Modelling Rainwater in Moscow

Ali J. Ramadhan[1*], *Soumik* Ray[2], *Mostafa* Abotaleb[3], *Hussein* Alkattan[3], *Garima* Tiwari[4], *Deepa* Rawat[5], *Pradeep* Mishra[6], *Shikha* Yadav[7], *Pushpika* Tiwari[8], Adelaja Oluwaseun Adebayo[9] and *Zainalabideen* Albadran[1]

[1]University of Alkafeel, Najaf, Iraq
[2]Centurion University of Technology and Management, Odisha, India
[3]South Ural State University, Chelyabinsk, Russia
[4]G. B. Pant University of Agriculture and Technology, Uttarakhand, India
[5]VCSG Uttarakhand University of Horticulture and Forestry, Uttarakhand, India
[6]Jawaharlal Nehru Krishi Vishwavidyalaya, Jabalpur, India
[7]University of Delhi, New Delhi, India
[8]Indian Institute of Forest Management, Bhopal, India
[9]Lagos State University, Lagos, Nigeria

Abstract. To model and forecast complex time series data, machine learning has become a major field. This machine learning study examined Moscow rainfall data's future performance. The dataset is split into 65% training and 35% test sets to build and validate the model. We compared these deep learning models using the Root Mean Square Error (RMSE) statistic. The LSTM model outperforms the BILSTM and GRU models in this data series. These three models forecast similarly. This information could aid the creation of a complete Moscow weather forecast book. This material would benefit policymakers and scholars. We also believe this study can be used to apply machine learning to complex time series data, transcending statistical approaches.

1 Introduction

Among one of the most compactly populated areas in the nation, the Moscow region ranks second among federal regions in terms of population density. Moscow is the biggest monocentric agglomeration in Europe, with a population of over 10 million people [1,2]. Government authorities are spread across Moscow and other areas in the oblast, therefore there is no administrative hub. The 2010 census counted 7,095,120 people; the 2015 census counted 7,231,068. The city is at 55 degrees, 45 minutes, and 7 seconds North and 37 degrees, 36 minutes, and 56 seconds East. Regional industrialization is high [3]. Moscow has a humid continental climate with short, warm summers and long, cold winters. The average temperature is 3.5°C (38.3°F) to 5.5°C (41.9°F). The coldest months are January and February. Western temperatures average -9°C (16 °F), whereas eastern temperatures average -12°C (10 °F).

* Corresponding author: ali.j.r@alkafeel.edu.iq

The lowest recorded temperature is -54 °C (-65 °F) (2, 3). This study uses a deep learning Long Short-Term Memory (LSTM) model and Google Earth Pro data from 1901 to 2019 to predict precipitation patterns. The current study found Moscow precipitation increased. Time series data prediction depends on the sequence being modeled and may have spatial dependencies across dimensions. Rainfall prediction has historically used physical and conceptual models [4]. Physical models can make near-daily forecasts and predictions, whereas conceptual models are better for daily time-scale phenomena. Conceptual and physical models that use temporal scales cannot accurately predict precipitation and runoff, especially when spatial resolution is low [5,6]. Additionally, these models require specific physical features that limit their capacity to reliably forecast sequence data with uncertain or limited quasi-periodic dynamics [7-10]. Rainfall predictions used LSTM, the closest statistical method to deep learning. According to Poornima and Pushpalatha [11], statistical methods are only accurate for time series data with non-uniform trends within and between seasons. A data analysis method could improve precipitation forecasts. The ability to estimate string data patterns and dynamics without parameters is responsible for this skill. It also accounts for observation randomness and system noise. Deep learning is preferred for examining precipitation time-series trends (Kratzert 2018). [12] noted that [13] suggested utilising LSTM for rainfall forecasting because to its higher performance in long-term analysis. The RBF network and empirical decomposition models are used in other precipitation forecasting studies, such as [14] study. The samples showed that the approaches failed to include seasonal degradation and cyclical oscillations in rainfall and runoff time sequence data. [15] developed an artificial neural network (ANN) training software that uses genetic algorithms to improve Kentucky River watershed flow-volume estimations. [16] studied India's sparse Malaprabha River Basin. To precisely mimic rainfall and runoff intensity, they used a modular neural network. [17] obtained better results than [18]. Internal repetition allows the LSTM technique to dynamically integrate past learning experiences. [19] found that LSTM is computationally more dependable and architecturally more coherent than feed-forward neural networks without internal states. LSTM can accurately replicate and imitate chaotic series by demonstrating time series data fundamental properties. LSTM is good for long time delays and low- and high-frequency time-series signals. Long Short-Term Memory (LSTM) can learn and capture time-series data better than ANN and statistical models like ARMA and ARIMA. This makes LSTM beneficial for assessing different time-series data [20,21].

LSTM is used to analyse and predict precipitation trends in Moscow, where meteorological and hydrological monitoring networks are lacking. To understand a model's significance in this setting. Due to a lack of meteorological data, LSTM benefits in precipitation applications have not been compared in the literature. Google Earth Pro meteorological data is used to analyse and predict Moscow rainfall patterns in this study.

2 Materials and Methods

A Recurrent Neural Network (RNN) uses the output from the previous step as input for the next phase. In classic neural networks, inputs and outputs are usually independent. However, earlier words are critical when predicting the next word in a sentence, so they must be preserved. Thus, the RNN was created to solve this problem using a Hidden Layer. RNNs are most known for their Hidden state, which stores sequence data. Recurrent Neural Networks (RNNs) have "memory" that stores previous calculations. It uses comparable parameters for each input because it produces the output by performing the same operation on all inputs or hidden layers. This simplifies parameters compared to other neural networks.

LSTM Model: Recurrent neural networks like LSTM can learn sequence prediction order dependency. The LSTM model network has had a big impact and can learn from sequential data. LSTMs can capture long-range dependencies, making them more precise than RNNs.

The backward flow of mistakes in the RNN architecture is complicated by back-propagation. The back-propagation approach uses the chain rule to combine functions. Finding the local loss function minimum with Stochastic Gradient Descent (SGD) is effective. After getting the gradients from the computational network nodes, we may calculate the gradients of nodes with a beginning point by finding its gradient and moving in the opposite way. Finding a local minimum is most efficient with this strategy. LSTM memory cells store and transmit unit outputs over several time steps. The LSTM memory cell uses cell states to store information over time. It also has a forgetting gate, input gate, and output gate to manage information flow between time intervals. This study predicted background radiation using time-series rainfall data and LSTM neural networks. The vanishing gradient problem arises when trying to explain recurrent neural network arrangement's persistent reliance. Initial stages have less impact as the input sequence lengthens, leading gradients for the first few input points to evaporate and reach zero. The LSTM activation function is the identity function, hence its output equals its input. The derivative of this function is always 1.0 because the LSTM is recurrent. Thus, backpropagated gradients remain constant. Nodes output according to their activation functions. LSTM networks use sigmoid and hyperbolic tangent activation functions. Current LSTM architecture uses sigmoid function for forget gate and input gate and tanh function for candidate vector that updates cell state vector.

BiLTSM Model: The hybrid BiLSTM model combines the strengths of the BiRNN and LSTM models. Information is sent forward and backward using the BiLSTM model. BiLSTM is a bidirectional neural network that captures past and future data. Among other parallel processing models is BiLSTM. The first step is encoding the wp and wq queries using two LSTM models.

GRU Model: Standard recurrent neural networks are improved by GRUs. GRU employs update and reset gates to tackle the conventional RNN vanishing gradient problem. The output information is determined by these two vectors. They are unique because they can be trained to retain old data without washing it or removing irrelevant data from the forecast.

3 Results and Discussion

From table 1, we found that; from January 1901 to December 2019, average rainfall increased from 1.30mm to 159.40mm. A Kurtosis value (3.52) indicates that data follows a mesokurtic distribution which shows the dataset has heavier tails than a normal distribution (more in the tails). Followed by a positive value of skewness (0.87) which is between more than 0.5, the distribution is approximately non-symmetric. These descriptive statistics show asymmetry in data set. This is the main reason that we have to choose neural network techniques.

Table 1. Descriptive statistics about the average monthly rainfall in Moscow from January 1901 to December 2019

	Minimum	Maximum	kurtosis	skewness	Standard deviation
Moscow	1.30	159.40	3.525279	0.8725834	27.25184

From table 2, we came to know about the accuracy of training data, which is 65% of the data set of monthly average rainfall in Moscow. For accuracy measurement, different statistics are given like; Mean Error (ME), Room mean square error (RMSE), mean absolute error (MAE), Mean absolute percentage error (MAPE), Mean percentage error (MPE) and Auto correlation function (ACF).

Table 2. Accuracy of training data 65% of the data set of monthly average rainfall in Moscow

Model	ME	RMSE	MAE	MPE	MAPE	ACF1
LSTM	0.104294	23.259798	18.61297	-24.963645	47.358616	-0.011684
BI-LSTM	1.608855	24.642836	19.637613	-24.24154	49.081123	0.041935
GRU	1.993111	24.696203	19.640396	-23.514741	48.790031	0.069637

From Table 3, it is clear that the LSTM model performed well for the given data set. LSTM networks were found to have more predictive skills after analyzing the data. Thevalues of ME, RMSE, MAPE, MAE and the MPE for LTSM are 0.104, 23.25, 47.35, 18.61 and -24.96. The value of all the errors for the training data set is lower in the case of LSTM model. For BiLTSM and GRU all the values are higher than LSTM. It shows the efficiency of the LTSM model. The value of ACF is -0.011 for LSTM, 0.041 for BiLTSM and 0.069 for GRU. The negative value of ACF showsa violation of independence. From table 3, we see the accuracy of test data i.e. 35% of the data set. For this we used RMSE and we see that the value of RMSE for the LSTM is 26.29, for BiLTSM are 27.70 and for GRU, it is 27.67.

Table 3. Accuracy of testing data 35% of the data set of monthly average rainfall in Moscow

Model	RMSE
LSTM	26.299781
BI-LSTM	27.700437
GRU	27.674776

From table 4, we conclude that the LTSM is perfect for test data as well as trained data. The accuracy of LSTM model is veryhigh because the values of the accuracy criteria (ME, RMSE, MAE, MPE, MAPE and ACF1)are lower than the values of the accuracy criteria of the BiLTSM and GRU model. So we outperform the forecasting of rainfall in Moscow by LSTM model.

Table 4. Forecasting monthly average rainwater in Moscow

Month	2020	2021	2022
January	44.973110	49.349667	50.016087
February	45.687706	48.818470	49.829872
March	46.940685	48.726063	50.206001
April	48.281284	49.677433	50.815350
May	50.242065	51.326649	51.298695
June	53.373997	52.700985	51.568611
July	56.263359	53.088612	51.759712
August	56.956448	52.696297	51.916859
September	54.769272	52.232498	51.902760
October	51.942791	51.900421	51.611423
November	50.343182	51.432720	51.157558
December	49.758602	50.699108	50.793095

In Table 4, monthly average rainfall is predicted by LSTM model. As the LSTM model is the best among the different RNN models,so, table 4 gives the best forecasted values for the present study. Forecasting is done for 36 consecutive months i.e. January 2020 to December 2022. Figures 1-3 depicts the trend of the forecast dataalong with prediction of training and testing data using all the three models LSTM, BiLTSM and GRU, models.

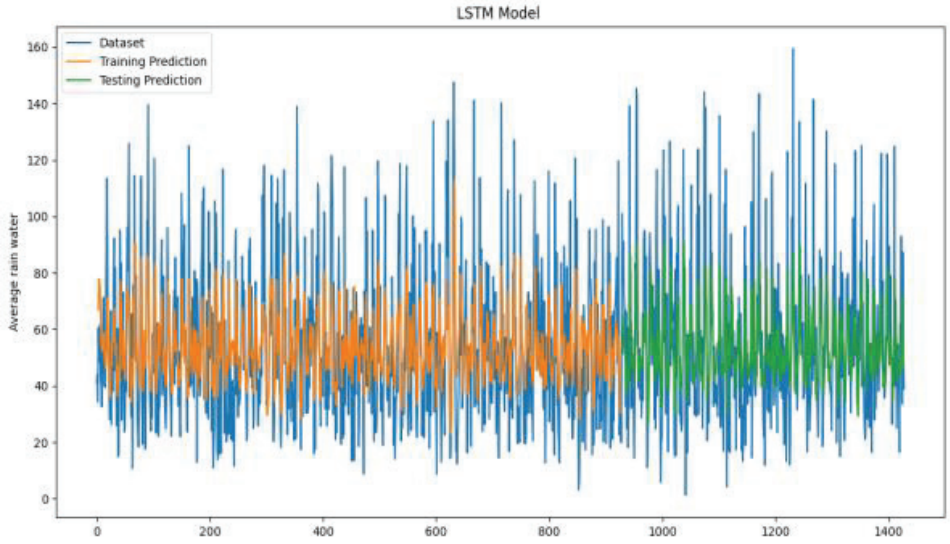

Fig. 1. LSTM model

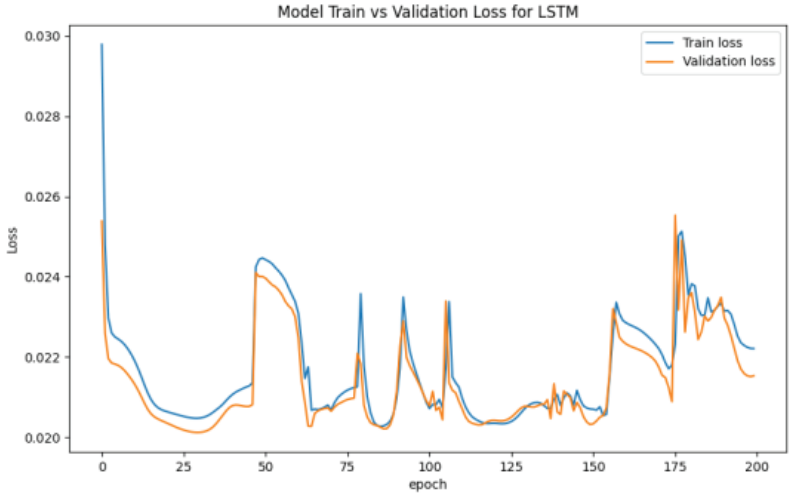

Fig. 2. Model Train vs Validation loss for LSTM

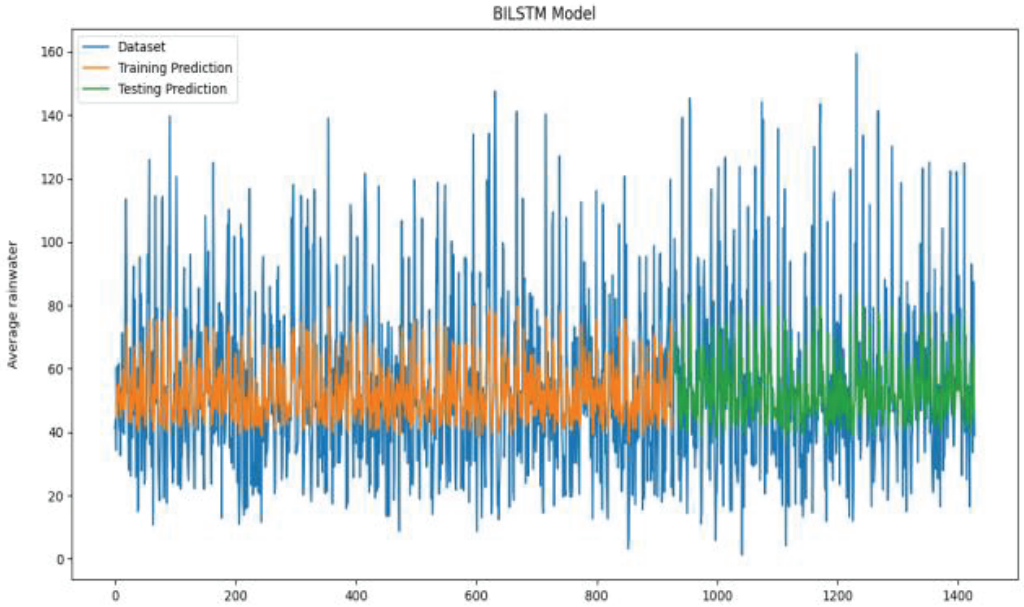

Fig. 3. BiLTSM model

Figures 4-6 gives a comparative picture between train loss and validation loss using LSTM,BiLTSM and GRU models. Algorithms for the complete methodology for all the used models viz LSTM, BiLTSM and GRU are shown in Fig 7.

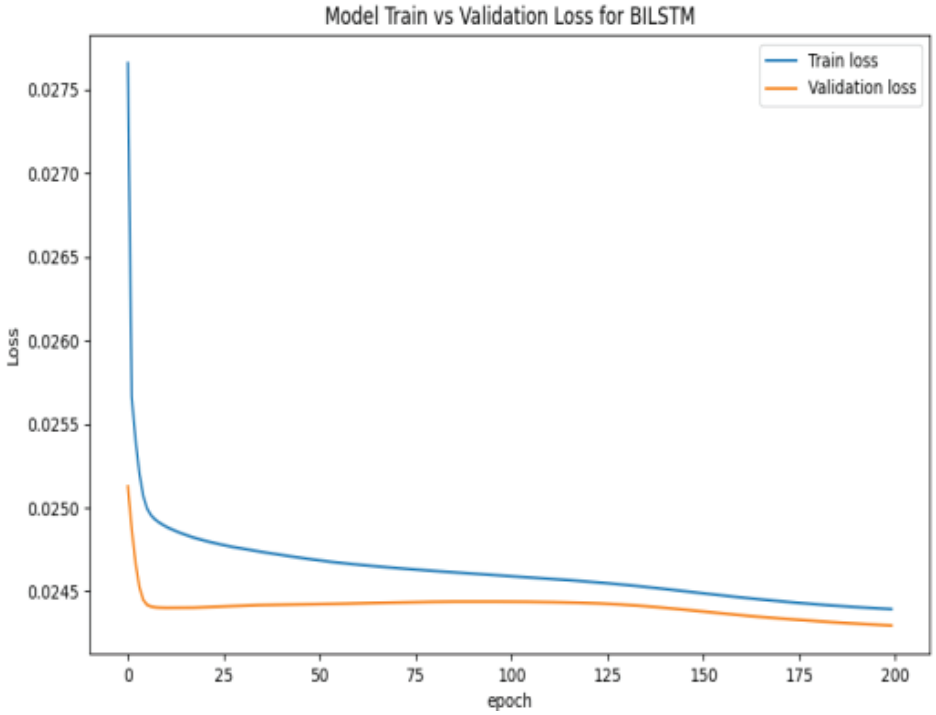

Fig. 4. Model Train vs Validation loss for BiLSTM

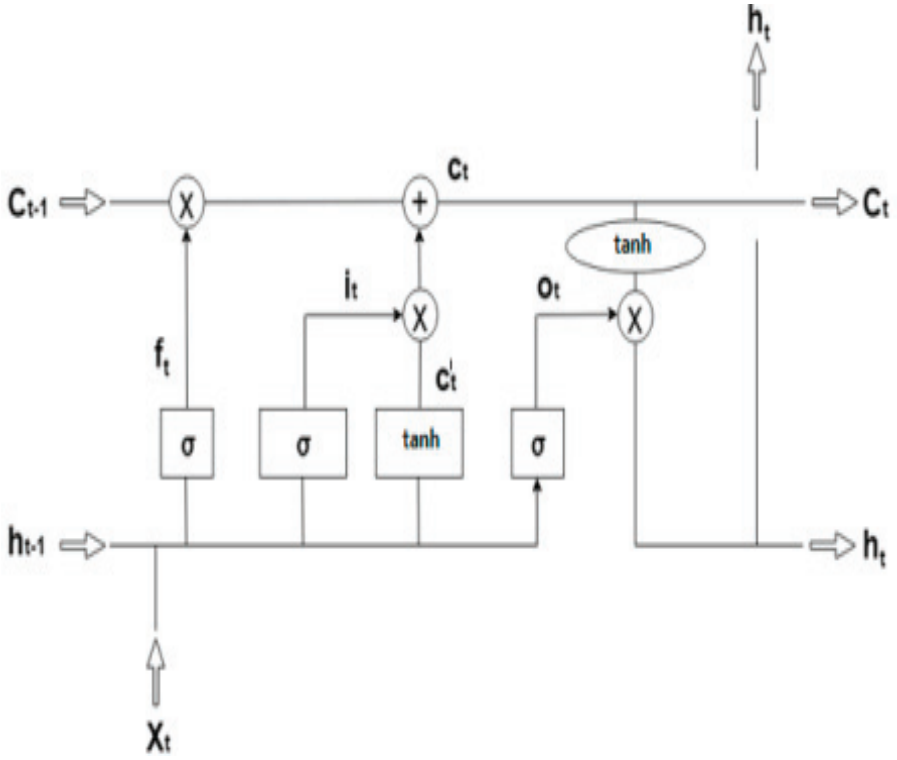

Fig. 5. The basic structural representation of LSTM network

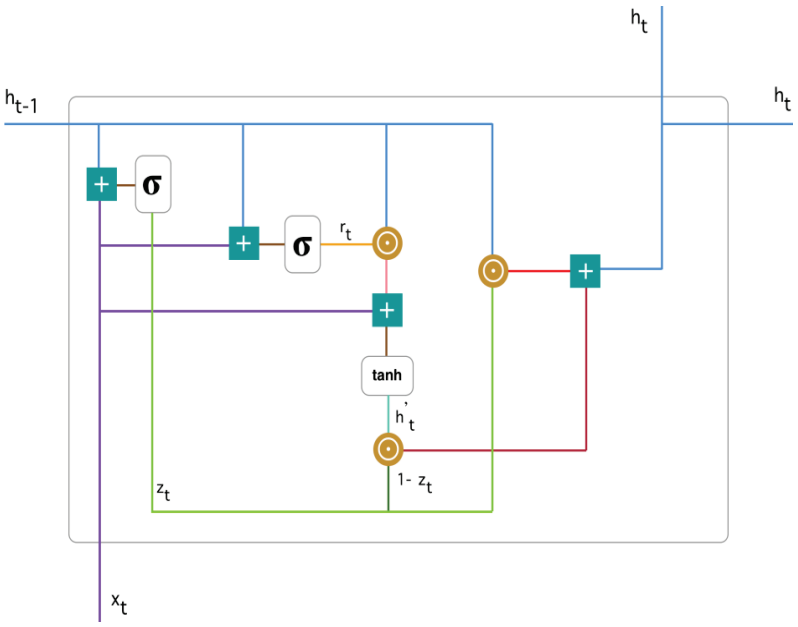

Fig. 6. Long short-term memory (LSTM) Neural.

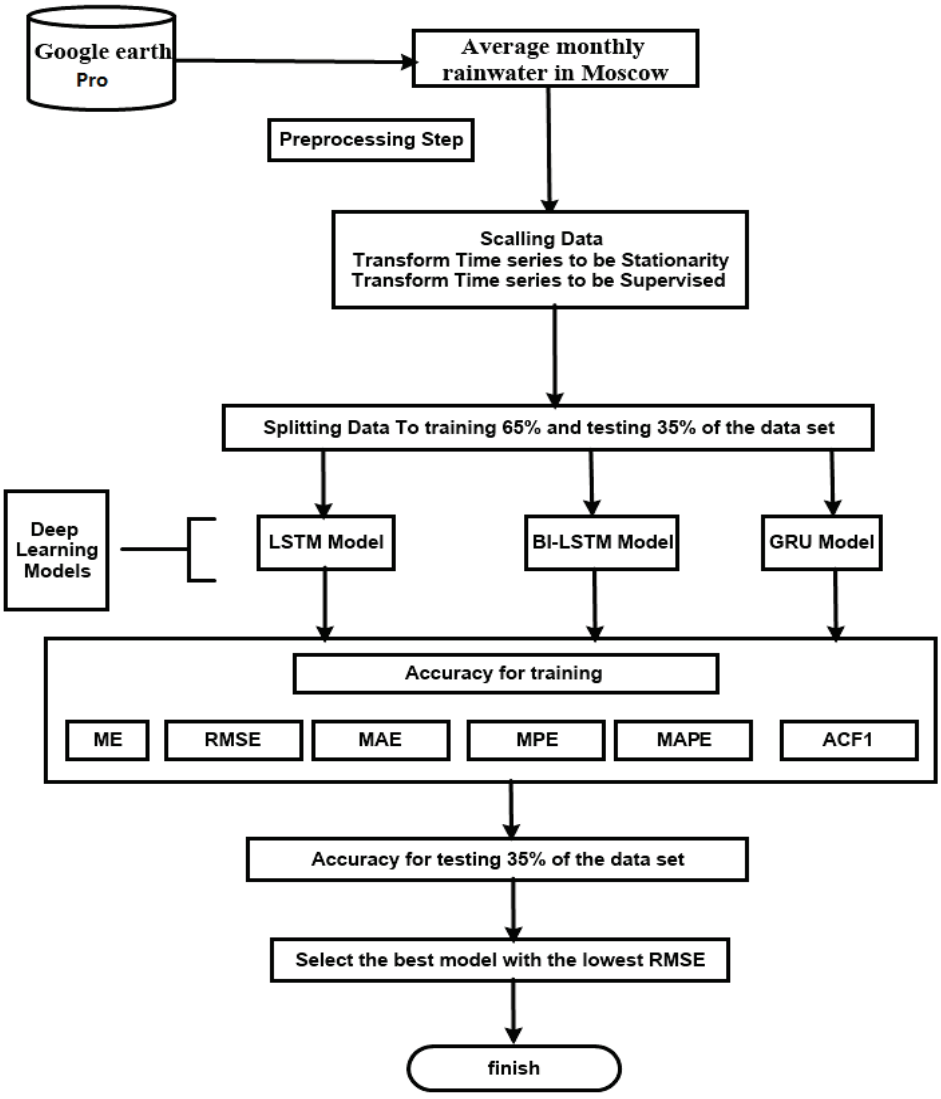

Fig. 7. Algorithm for the complete methodology for all the used models viz. LSTM, BiLTSM, and GRU

In this work we show that the rainfall data of Moscow can be used to predict future behavior with machine learning technique. We obtained a good result using the deep RNN approach. Three deep learning approaches *i.e.* LSTM, BiLSTM and GRU models are used to predict the forecast of rainfall in Moscow, and tried to find out the best model based on different goodness of fit in both the training and the testing data set. The LSTM model gave a batter approach over the other two models. For testing data set 5.06% RMSE improvement was observed int he LSTM than the BiLSTM and 4.968% RMSE improvement than the GRU model. Allover, these results confirmed a strong acceptance of a deep learning approach to estimating the weather forecast.

4 Conclusion

In the projection model, every model shows a different error rate.In the present investigation; there are different measures of errors, viz. ME, RMSE, MAE, MAPE, MPE and ACF1 for the training data.For testing data, there is RMSE for checking the accuracy and efficiency of the models. The best model for forecasting monthly rainfall in Moscow is the LSTM model. This model shows minimum errors in all the aspects. For training data set, LSTM model has a minimum RMSE that is 23.25, minimum MAPE, MAE, MPE and ME *i.e*.47.35, 18.61, -24.96 and 0.10 respectively. For the testing data set, LTSM has a minimum RMSE i.e. 26.29. So, for both the process,the LTSM model shows the best results. Although BiLTSM and GRU models perform well but they are not best just because of their highest measures of accuracy. The present investigation helps to study the irregular past trends and gives the best prediction model.

References

1. Beven, K. J., Cloke, H. L., 2012, “Comment on ‘hyper resolution global land surface modeling: Meeting a grand challenge for monitoring Earth’s terrestrial water’ by Eric F. Wood et al.,” Water Resour. Res., **48**, 1, 1–10.
2. Ehsan khodadadi, S. K. Towfek, Hussein Alkattan. (2023). Brain Tumor Classification Using Convolutional Neural Network and Feature Extraction. Fusion: Practice and Applications, **13**(2), 34-41.
3. Corzo, G., Solomatine, D., 2007, “Baseflow separation techniques for modular artificial neural network modelling in flow forecasting”, Hydrol. Sci. J., **52**, 3, 491–507.
4. Al-Nuaimi, B. T., Al-Mahdawi, H. K., Albadran, Z., Alkattan, H., Abotaleb, M., & El-kenawy, E. S. M. (2023). Solving of the inverse boundary value problem for the heat conduction equation in two intervals of time. Algorithms, **16**(1), 33.
5. Gursoy, O., Engin, S. N., 2019, “A wavelet neural network approach to predict daily river discharge using meteorological data”, Meas. Control (United Kingdom), **52**, 5–6, 599–607.
6. Akbari, E., Mollajafari, M., Al-Khafaji, H. M. R., Alkattan, H., Abotaleb, M., Eslami, M., & Palani, S. (2022). Improved salp swarm optimization algorithm for damping controller design for multimachine power system. IEEE Access, **10**, 82910-82922.
7. Jain, A., Srinivasulu, S., 2006, “Integrated approach to model decomposed flow hydrograph using artificial neural network and conceptual techniques”, J. Hydrol., **317**, 3–4, 291–306.
8. Kratzert, F., Klotz, D., Brenner, C., Schulz, K., Herrnegger, M., 2018, “Rainfall-Runoff modelling using Long-Short-Term-Memory (LSTM) networks”, 1–26.
9. Al-Mahdawi, H. K., Albadran, Z., Alkattan, H., Abotaleb, M., Alakkari, K., & Ramadhan, A. J. (2023, December). Using the inverse Cauchy problem of the Laplace equation for wave propagation to implement a numerical regularization homotopy method. AIP Conference Proceedings (Vol. **2977**, No. 1). AIP Publishing.

10. Lindsay, G. W., 2021, "Convolutional neural networks as a model of the visual system: Past, present, and future", J. Cogn. Neurosci., **33**, 10, 2017–2031.
11. Liu, M., et al., 2020, "The applicability of lstm-knn model for real-time flood forecasting in different climate zones in China", Water (Switzerland),**12**, 2, 1–21.
12. Parkes, B. L., Wetterhall, F., Pappenberger, F., He, Y., Malamud, B. D., Cloke H. L., 2013, "Assessment of a 1-hour gridded precipitation dataset to drive a hydrological model: A case study of the summer 2007 floods in the upper severn, UK", Hydrol. Res., **44**, 1, 89–105.
13. Poornima, S., Pushpalatha, M., 2019, "Prediction of rainfall using intensified LSTM based recurrent Neural Network with Weighted Linear Units', Atmosphere (Basel)., 10,11.
14. Sang, Y. F., 2013, "A review on the applications of wavelet transform in hydrology time series analysis", Atmos. Res., **122**, 8–15.
15. The Best Time to Visit Chelyabinsk, Russia for Weather, Safety, & Tourism, champion Traveler, https://trek.zone/en/russia/places/18580/chelyabinsk.
16. Weather and Topography of Chelyabinsk (The weather year-round anywhere on earth), Weather Spark, https://weatherspark.com/y/106113/Average-Weather-in-Chelyabinsk-Russia-Year-Round.
17. Wesemann, J., Herrnegger, M.,Schulz, K., 2018, "Erratum to: Hydrological modelling in the anthroposphere: predicting local runoff in a heavily modified high-alpine catchment," J. Mt. Sci., p. 1.
18. Wong, K. W., Wong, P. M., Gedeon, T. D., Fung, C. C., 2003, "Rainfall prediction model using soft computing technique", Soft Comput., **7**, 6, 434–438.
19. Xiang, Z., Yan, J., Demir, I., 2020, "A Rainfall-Runoff Model With LSTM-Based Sequence-to-Sequence Learning", Water Resour. Res., **56**, 1.
20. Zhang, B., Govindaraju, R. S., 2000, "Prediction of watershed runoff using Bayesian concepts and modular neural networks", Water Resour. Res., **36**, 3, 753–762.
21. Zolotokrylin, A. N., Vinogradova, V. V., Titkova, T. B., Cherenkova, E. A., Bokuchava, D. D., Sokolov, I. A., Vinogradov, A. V., Babina, E. D., 2018,"Impact of climate changes on population vital activities in Russia in the early 21stcentury". IOP Conf Ser: Earth and Environ Sci, **107**:012045. https://doi.org/10.1088/1755-1315/107/1/012045