Received: December 1, 2021. Revised: January 25, 2022.

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

Rainfall Prediction Based on Spatial Attention Layer: A Case Study Analysis

Subha Sree Mallela[1]* **Sravan Kumar Jonnalagadda[1]**

[1]*Department of Computer Science and Engineering, KKR & KSR Institute of Technology and Sciences, Guntur, India*
* Corresponding author's Email: subhasree87@gmail.com

Abstract: Indian summer Monsoon Rainfall (ISMR) prediction models help in the field of agriculture, water management and disaster prediction at an early phase to avoid unnecessary loss of lives as well as economy. In this research, Spatial Attention Layer - Long Short Term Memory (SAL-LSTM) method is used to solve a univariate time series forecasting problems and it is used in the ISMR prediction at an early phase. Spatial Attention layer consists of two phase namely hard attention and soft attention for selection of the input data and weighting the value. The sigmoid function is applied in the hard attention for the selection of the input data and softmax function is applied in the soft attention for weighting the input value. Hadmard product is applied in the attention layer to handle the spatial data and Rectified Linear Unit (ReLU) activation function is applied in attention layer to handle the temporal data. The data is collected from an Indian Institute of Tropical Meteorology (IITM) of Kerala, to test the model performance in Rainfall prediction. The Min-Max Normalization technique is applied to solve the problem of missing data in the IITM dataset. Melbourne rainfall dataset were used to evaluate the performance of SAL-LSTM method and compared with LSTM method. The experimental outcome explains that the proposed SAL-LSTM method has higher performance in terms of achieving lower Mean Absolute Error (MAE) values compared to existing LSTM, Artificial Neural Network (ANN) and Extreme Learning Machine (ELM) methods in IITM and Melbourne datasets. The MAE of the proposed SAL-LSTM is achieved as 4.213 and existing LSTM method achieves 5.653 MAE. The proposed spatial attention layer has 0.12 MAE and existing attention layer with LSTM has 0.17 MAE in Melbourne dataset.

Keywords: Hidden layers, Indian institute of tropical meteorology, Indian summer monsoon rainfall, Long short term memory, Min-max normalization.

1. Introduction

Rainfall is one of the important factors in hydrological model that affects the quantity and quality of local water resources. The small changes of rainfall levels may cause severe flooding or drought, which directly affects food production and even the economic activities [1-3]. The accurate prediction of rainfall (precipitation) is challenging due to the complexity of meteorological phenomena. Rainfall is caused by a variety of meteorological conditions that are computed by using a mathematical model and the duration between the rainfalls is predicted from the starting month of the rainy season and the end of the rainy season are also calculated [4, 5]. Monthly rainfall values provide more accurate an intra-year rainfall distribution than seasonal rainfall values [6, 7]. The frequency of occurrence and intensity of heavy rainfall event over the Indian subcontinent has been increasing in the past few decades. Recently, a heavy rainfall has occurred over the Indian sub-continent that caused huge damage, loss to life, property and such losses can be significantly reduced by providing more accurate prediction of these events in advance [8-10]. Several methods have been proposed by various researchers for modelling rainfall data in India [11]. The regression problems in the most of the research are challenging in machine learning because, the predictions of the target variables are critical with respect to the intelligent system for a specified application [12, 13]. The Artificial Neural Network (ANN) provides the higher efficiency in predicting the rainfall and has been used in the many existing methods [14-17]. Wavelet analysis has been applied by various researchers in the field of hydrological modelling, as the wavelet analysis is a useful for time frequency characterization of a time series and wavelet based conjunction models for hydrological modelling shows more significance in this research. There are a large number of the conceptual basis of wavelet and NNs used for forecasting rainfall [18, 19] However, the accurate forecast of ISMR is desirable in the proposed LSTM, even if the difficult tasks are interrupted in the scientific community. Among different scientific approaches, numerical modelling with empirical and statistical approaches are popular that are used widely in the recent time for monsoon forecast, whereas the soft computing still not gained the attention [20]. Existing LSTM method has the limitation of lower efficiency in handling the spatio-temporal data. In this research, SAL-LSTM model is proposed to effectively analysis the data and improve the performance of rainfall prediction. The proposed SAL-LSTM model is evaluated in the two datasets such as Kerala dataset and Melbourne dataset. The new features of the proposed SAL-LSTM model are explained as follows:

1. Attention layer of two phase namely hard attention and soft attention is applied in LSTM model. Hard attention phase selects the input data for LSTM data based on the sigmoid function. Soft attention involves in weighting of features based on the softmax function.
2. Hadmard product is applied in the attention layer to handle the spatial data and ReLU activation function is applied to handle the temporal data in LSTM.
3. The proposed method attention method preserves the spatial temporal information that helps to improves the performance of prediction.
4. Adam optimization method is applied to train the LSTM model with optimal parameter settings. Adam optimization method selects the optimal learning rate for the LSTM model.

The SAL-LSTM is applied to the normalized data for rainfall prediction that are used for evaluating the performance measures for rainfall prediction. The rainfall prediction is carried out for the two monsoon seasons; Summer Monsoon which occurs in the month of June July August September (JJAS) and post-monsoon which occurs in the month of October November December (OND).

The organization of the paper is given as follows: Literature Survey is presented in Section 2 and the proposed methodology is explained in Section 3. The experimental results of the proposed method are tabulated in Section 4 and the conclusion of this paper is given in Section 5.

2. Literature survey

Rainfall prediction is the difficult process due to the presence of many uncertain factors and those factors are required to be analyzed effectively. Recent researches related to the rainfall prediction in India are reviewed in this section.

Dash [21] applied K-Nearest Neighbor (KNN), ANN and Extreme Learning Machine (ELM) for rainfall prediction based on seasonal forecasting. The time series data were collected from the IITM, Pune for Kerala division. The Data Normalization was carried out in the research for pre-processing step based on Min-Max Normalization. The ELM obtained performance at a higher rate in rainfall prediction when compared with KNN and ANN classification method. The error values of the three classifiers were considerably low in the monsoon prediction. ELM has the effective performance in small dataset and ELM has the low effective in handling the large dataset. ELM doesn’t effectively analysis the relation between the input and output data especially it has lower efficiency in time series dataset.

Dash [22] developed the classifiers of Logistic Regression (LR), ANN and ELM with considered a global sea surface temperature. The Principal Component Analysis (PCA) is used for the dimensionality reduction of the data sets and this improves the predictive ability of machine learning. The research shows that global sea surface temperature is one of the important parameters in the rainfall prediction. In order to perform data normalization, min-max normalization technique was performed for the collected dataset. The weighting factor of ANN has lower efficiency in handling the time series data and lower efficiency in rainfall prediction.

Vathsalaand Koolagudi [23] developed a hybrid model to predict the Indian monsoon rainfall and applied an algorithm to find the suitable model for processing the dataset. The data are collected from the ISMR and 36 factors from the datasets are used for the prediction technique. The generation based association rule mining technique is used for the feature selection method. The cluster membership is used for the dimensional reduction of the dataset and applied logistic function to predict the rainfall and it reduces the factors from 25 to 8 for the logistic to predict the rainfall. The results obtained from the experiment shows that the significant performance of error values is achieved by developed method. The logistic function has the disadvantages of assuming the linearity between the dependent and independent variable that provides the lower efficiency in the prediction.

Devi [24] developed different neural network model such as Backpropagation neural network, Cascade Forward Back Propagation Neural Network (CFBPNN), distributed time delay neural network and non-linear auto regression exogenous network. The developed method is tested on the two datasets one containing the data of Nilgiris and other has some gauge station. The three classifiers have the considerable performance in the rainfall prediction. The study shows that the Levenberg Marquardt is the effective weight update technique compared to the other gradient method. The method was failed to predict the high peak of the dataset and feature selection method is need to be applied to effective performance.

Pham [25] developed a hybrid method of Particle Swarm Optimization (PSO), Support Vector Machine (SVM) and ANN for the rainfall prediction. The meteorological data were collected and used as the input for the prediction model. The results obtained from the developed method shows good performance in rainfall prediction. The Monte Carlo method validated the performance measures from the developed method. The data were collected from the Global weather data for SWAT. The developed method has the highest performance in the rainfall prediction. However, for rainfall prediction for the transit period from a non-rain to a rainy day was failed to indicate the strengths or weaknesses of the models.

Pudashine [26] deep learning LSTM model architecture is applied for rainfall prediction in Melbourne dataset. The analysis shows that the LSTM method has the considerable performance in the rainfall prediction. The LSTM model has the higher performance compared to dense layer ANN, and Gated Recurrent Unit (GRU). The LSTM model has the limitation of lower efficiency in handling the spatio-temporal rainfall data. The weighting value of the LSTM model doesn't effectively analysis the relation of input features to the output.

Sangeetha and Prabha, [27] proposed Multi-Head Attention Layer (MHAL) with fine grid embedding in input sequence and tested with dropout layer. The information from both deep multi-layer is fused and applied as input to LSTM model. The single attention pass over value is applied to use weight value in multihead attention layer and weight value is applied in LSTM model for classification. The two heads output are applied in the dropout layer and integrated in concatenation layer that output is given in LSTM method. The glove and clove embedding methods involves in applying weight value for irrelevant features in classification.

Liu and Guo, [28] proposed attention mechanism, Bi-directional Long Short Term Memory (BiLSTM), and Convolution layer (AC-BiLSTM) for classification process. The convolution layer extracts high level phase representation from work embedding vectors and BiLSTM model to access context representation. Attention method is applied to provide different focus on information outputted from hidden layer of BiLSTM. The softmax layer is applied to classify the context information. The AC-BiLSTM captures both local and global features in the input sequence. The developed method applies weight value to words in the sentence to consider the sentiment and this tends to overfit the model.

Differ from existing method: The existing attention method involves in selects the features based on embedding vector and weight value for hidden layers. Existing attention method tends to lose the spatial temporal information in network. The proposed spatial attention layer involves in applying the hard and soft attention layer to preserve the spatial temporal information that are important for time series prediction. The proposed spatial attention layer performs hard attention and soft attention in input sequence. The hard attention decides whether to consider the input sequence for the classification and soft attention provides more flexible value to input. The input with weight value is applied to train the model and improve its learning performance.

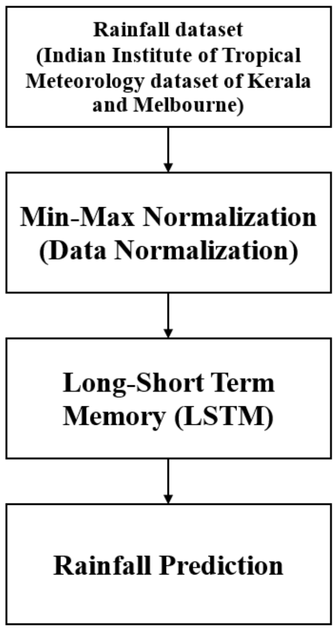

Figure. 1 The block diagram of the proposed LSTM method

Received: December 1, 2021. Revised: January 25, 2022.

3. Proposed methodology

Rainfall prediction is the challenging research due to the presence of uncertain factors in the classifiers. In this research, the dataset is collected from IITM, Kerala and LSTM classifier is proposed to increase the effectiveness of the rainfall prediction. The LSTM classifier is able to store the important information for the long term and effectiveness analysis the features. The block diagram of the proposed LSTM is shown in the Fig. 1.

3.1 Dataset

This research work uses time series data as the dataset which is collected from IITM, Kerala subdivision from the year 1871 to 2016 [21]. The monthly data having area weighted of the rainfall are used for this research. The Summer Monsoon (JJAS) and post-monsoon (OND) months mean time series were used for prediction that receives maximum rainfall for the respective years. Apart from Kerala dataset, Melbourne dataset [26] were used to evaluate the performance of the proposed SAL-LSTM method. Melbourne dataset consists of 7874 instances related to the rainfall data from December 2017 to November 2018.

3.2 Data normalization

The data normalization also refers to the data scaling which is a pre-processing step that is necessary before training phase for neural network initiation. The data normalization is performed using Min-Max normalization approach and the inputs of the neurons performs normalization as shown in the Eq. (1).

$$y^* = y_{min} + (y_{max} - y_{min}) \times \frac{x - x_{min}}{x_{max} - x_{min}} \quad (1)$$

From the Eq. (1), $y*$ is represented as scaled value , y_{min} is the minimum value y_{max} is the maximum value that lies in the range of 0 to 1 respectively. x is the original rainfall value need to be normalized. The x_{min} is the minimum value and $x_{max}$is the maximum value of rainfall data.

3.3 Long short term memory

The financial time series requires previous as well as the latest data for the prediction. The Self-feedback mechanism benefits from hidden layer which acts as an added advantage using Recurrent Neural Network (RNN) model in dealing with dependence problems at long term and there are no difficulties with respect to practical application.

The LSTM unit has a memory cell for storing the information and they generally updated with the help of 3 special gates such as input gate, forget gate and output gate [29, 30, 31]. The block diagram of LSTM network is shown in the Fig. 2.

At the time t the data x_t is given as input to the LSTM cell, Where the output of the LSTM cell is h_{t-1} at the previous state, the memory cell value is represented as c_t, the output of the LSTM cell is h_t, The LSTSM unit calculation process is divided into the following steps.

1. Calculate the value of the $\widetilde{c_t}$ which is the candidate memory cell, W_c is known as the weight matrix, the bias is represented as b_c. The value of $\widetilde{c_t}$ is calculated by using Eq. (2).

$$\widetilde{c_t} = \tanh(W_c . [h_{t-1}, x_t] + b_c) \quad (2)$$

2. The input gate i_t should be calculated, where the input gate controls the update of the current input data to the state value of the memory cell,W_i is the

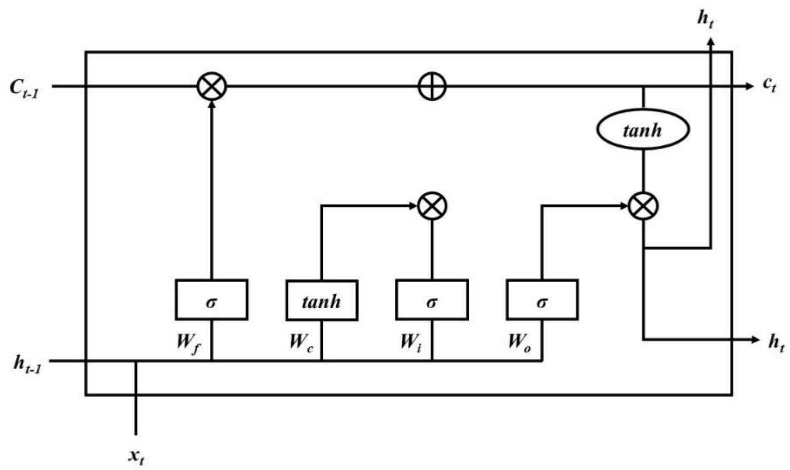

Figure. 2 The block diagram of LSTM structure

weight matrix, σ is sigmoid function, b_i is the bias. The value of i_t is calculated by using Eq. (3).

$$i_t = \sigma(W_i.[h_{t-1}, x_t] + b_i) \quad (3)$$

3. The value of the forget gate f_t is calculated that controls the forget and historical data is updated to the memory cell from the state value, b_f is the bias, W_f is the weight matrix. The value of f_t is calculated by using Eq. (4).

$$f_t = \sigma(W_f.[h_{t-1}, x_t] + b_f) \quad (4)$$

4. The current moment memory cell c_t, and the last LSTM unit state values $c_{t-1}$are calculated. The value of $c_t$is calculated by using Eq. (5).

$$c_t = f_t \times c_{t-1} + i_t * \widetilde{c_t} \quad (5)$$

From the Eq. (5), the memory cell updates based on the last cell state value of the candidate cell that are controlled by forget and input gate.

1. The value of output gate o_t value is calculated, where the output gate controls the state value of the memory cell output, b_o is the bias and W_o is the weight matrix.

$$o_t = \sigma(W_0.[h_{t-1}, x_t] + b_0) \quad (6)$$

2. At last, the output of LSTM unit $h_t$is calculated.

$$h_t = o_t \times \tanh(c_t) \quad (7)$$

The LSTM has 3 control gates, memory cell, read, reset and update according to the long-time availability of information. The sharing mechanism importance in LSTM is constructed based on internal parameters and obtains output dimensions that are controlled by setting the weight matrix dimensions. The LSTM establishes a long time delay between the input and feedback, because the internal state of the memory cell and gradients are exploded or disappears if it maintains the error flow continuously.

3.4 Spatial attention layer

Neural network is applied with the pre-processed input sequence to generate raw attention weights. The final output is generated to provide the softmax attention weights and raw input value. Attention models are classified into two kinds, hard attention and soft attention. Hard attention refers to One-hot selection of input data that means attention weight value can either 0 or 1. The soft attention refers to weight between 0 and 1 and weight selection value is more flexible. First, input features are assigned with dynamic feature weights (spatial weights). Then, each step of LSTM hidden layer state is fully used to dispatch the temporal attention weights of each time step. Input and output of LSTM is affected by spatial and temporal attention weight value. The spatial and temporal attention weight values is used to dynamically adjust the attention weights and improve the efficiency of LSTM cell. Adam algorithm is used to train the models; spatio-temporal model is explained as below.

3.4.1. Spatial attention operation

Considered a 2D spatio-temporal feature matrix$X \in R^{m \times k}$, where the number of features in a single time step is denoted as m and k is the number of time steps. Rainfall data of multiple stations are given as input to this model. The input feature matrix X can be divided into k m-dimension vector is shown in Eq. (8).

$$x_t = [f_1^t, f_2^t, \dots, f_m^t]_{m \times 1} \quad (8)$$

$$\alpha_t = SA(x_t) = [\alpha_1^t, \alpha_2^t, \dots, \alpha_m^t]_{m \times 1}^T \quad (9)$$

$$x_t' = \alpha_t \odot x_t = [\alpha_1^t f_1^t, \alpha_2^t f_2^t, \dots, \alpha_m^t f_m^t]_{m \times 1} \quad (10)$$

The input feature vector is activated by $sigmoid(x) = \frac{1}{1+e^{-x}}$ after the monolayer neurons calculation. The spatial attention weight α_t is calculated based on the normalization of $softmax(x_i) = \frac{e^{-x_i}}{\sum_{i=1}^{n} e^{-x_i}}$, as shown in Eq. (9). Softmax is usually used for normalization to make sure limited additivity of weights. Hadamard product is given in Eq. (10).

3.4.2. Temporal attention layer

LSTM cell is applied step by step for spatial-attention-weighted sequence data. The output sequence of hidden layer states is given in Eq. (11).

$$H = [h_1, h_2, \dots, h_k]_{k \times s} \quad (11)$$

$$\beta = TA(H) = [\beta_1, \beta_2, \dots, \beta_k]_{1 \times k} \quad (12)$$

$$h_{att} = \beta \otimes H = \sum_{i=1}^{k} \beta_i h_i,\ h_{att} \in R^{1 \times s} \quad (13)$$

$$p = O(h_{att}), p \in R^{1 \times n} \quad (14)$$

Temporal attention weight β is generated based on ReLU activation ($ReLU = \max(0, x)$) and softmax normalization as given in Eq. (12). The matrix product $\otimes$ is given in Eq. (13) and the final prediction p is generated by output layer without any activation layer, as shown in Eq. (14).

4. Experimental results

Accurate prediction of rainfall helps in agriculture, water management, and disaster management at an early stage so that the future consequences can be overcome. The existing method involves applying classifiers such as ANN, KNN and ELM for the rainfall prediction. In this research, LSTM is proposed for the rainfall prediction to increase the efficiency of the method in terms of performance measures such as MAE, Root Mean Square Error (RMSE). The data are collected from the IITM of Kerala and Min-Max Normalization is applied to solve the problem of missing data. The rainfall data are analyzed in the two categories based on the monsoon seasons such as October November December (OND) and June July August September (JJAS). The dataset was divided into two parts, data collected from 1876 to 2010 are considered as training data and the data collected from 2011 to 2016 are used as the testing data.

Experimental design: The developed LSTM method is tested on the system consists of Intel i7 processor with 16 GB of RAM and 500 GB hard disk and both proposed LSTM and existing methods were implemented and compared in the same data and same environment.

Metrics used: In this research, the error values such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Mean Absolute Scaled Error (MASE) were used to analyse the performance of the developed and existing ELM, ANN methods. The formula for measure MAE, RMSE and MASE are expressed in the Eq. (15) to (17) respectively.

$$MAE(\%) = \frac{100}{x}\sum_{t=1}^{x}\left|\frac{Obs_t - Prd_t}{Obs_t}\right| \tag{15}$$

$$RMSE(\%) = \sqrt{\frac{100}{x}\sum_{t=1}^{x}\left(\frac{Obs_t - Prd_t}{Obs_t}\right)^2} \tag{16}$$

$$MASE = \frac{1}{x}\sum_{t=1}^{x}\left(\frac{|Obs_t - Prd_t|}{\frac{1}{x-1}\sum_{i=1}^{2}|Obs_i - Obs_{i-1}|}\right) \tag{17}$$

where Obs denotes the observed value, Prd denotes the predicted value for the year t and x denotes the total number of years.

Parameter Settings: The LSTM model has 1 input layer, 3 hidden layer and 1 output layer. The learning rate of LSTM is set as 0.01, dropout rate is 0.01, epochs is 20, and adam optimizer is used.

4.1 Impact on hidden nodes

The proposed LSTM and existing methods such as ANN and ELM have been evaluated in the rainfall prediction with various number of hidden nodes. The error value is measured for the various number of hidden nodes and is compared with the existing methods. The proposed SAL-LSTM method is evaluated and compared with LSTM, ANN and ELM in JJAS rainfall data, as shown in Table 1. The analysis shows that the proposed SAL-LSTM method has lower error value compared to ANN and ELM method. The proposed SAL-LSTM method has the advantage of storing the important information for long term and handle the spatio-temporal data. This helps to improve the performance of the SAL-LSTM classifier. The SAL-LSTM method has RMSE of 2.105 for 15 hidden nodes and ELM has 2.827 RMSE.

The comparison of proposed LSTM and existing methods such as ANN and ELM in JJAS Rainfall prediction is presented in Fig. 3. The error metrics such as MAE and RMSE were measured in the prediction technique. The number of hidden layers are varying from 10, 15 and 20 for the existing ANN, ELM, LSTM and proposed SAL-LSTM method. The performance metrics show that the proposed SAL-LSTM method has the lowest error value in the rainfall prediction due to the capacity of the model to store the important feature for classification in the long term. The error value of the proposed SAL-LSTM method has achieved MAE as 4.213 in the 20 hidden layers while LSTM method achieved as 5.653 error value.

The analysis shows that the 15 hidden layers has the higher performance compared to 10 and 20 hidden layer and as the more hidden layers tend to

Table 1. JJAS rainfall prediction error value

Techniques	No of Hidden Nodes	MAE	RMSE
ANN [21]	10	10.191	12.911
	15	7.999	9.297
	20	6.189	6.753
ELM [21]	10	5.095	6.886
	15	3.075	3.809
	20	7.279	9.827
LSTM	10	4.624	4.256
	15	2.168	2.827
	20	5.653	7.863
SAL-LSTM	10	3.126	3.576
	15	1.724	2.105
	20	4.213	5.764

Received: December 1, 2021. Revised: January 25, 2022.

store irrelevant features that affects the performance of the model.

The comparison of the proposed SAL-LSTM and existing methods in rainfall prediction in the monsoon of OND are shown in the Fig. 4. The comparison results show that the proposed SAL-LSTM showed higher performance than existing LSTM, ELM and ANN methods. The proposed SAL-LSTM results with lower error value in the rainfall prediction due to its ability to hold the important information for the long term.

The existing ELM method has the considerable performance in the rainfall prediction. The proposed SAL-LSTM and existing method are analyzed in the 10, 15 and 20 number of hidden layers in the prediction. The 15 hidden layers have the highest performance ELM and LSTM. The proposed SAL-LSTM is evaluated and compared with LSTM, ANN and ELM on OND method, as shown in the Table 2. The analysis shows that the proposed SAL-LSTM method has lower error value compared to LSTM, ANN and ELM method.

The proposed LSTM method has the advantage of storing the important features for long term and perform the classification. This increases the performance of the proposed SAL-LSTM method in rainfall prediction. The analysis shows that the proposed SAL-LSTM method has RMSE of 3.729 in 15 hidden nodes and LSTM method has 4.267 RMSE in 15 hidden nodes.

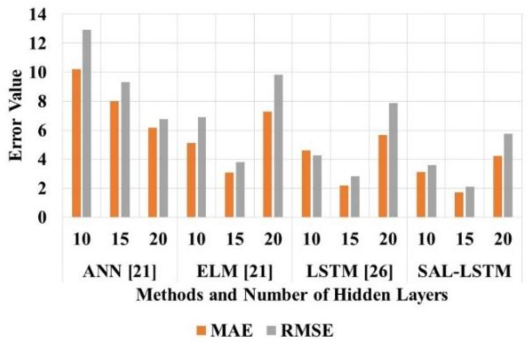

Figure. 3 JJAS rainfall prediction error value

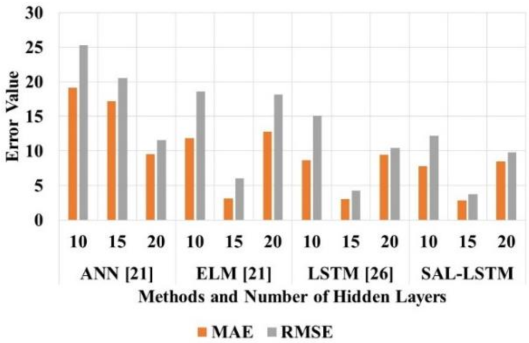

Figure. 4 Analysis of OND for existing and proposed method

Table 2. The performance analysis of proposed SAL-LSTM on OND

Techniques	No of Hidden Nodes	MAE	RMSE
ANN [21]	10	19.096	25.282
	15	17.225	20.543
	20	9.514	11.542
ELM [21]	10	11.832	18.59
	15	3.149	6
	20	12.781	18.147
LSTM	10	8.625	15.126
	15	3.024	4.267
	20	9.426	10.424
SAL-LSTM	10	7.768	12.155
	15	2.817	3.729
	20	8.457	9.776

4.2 Prediction performance

The prediction performance of the proposed LSTM is analyzed in the sub-section to investigate its effectiveness. The error value such as MAE, RMSE, and MASE are measured from the proposed SAL-LSTM prediction method.

The performance measures obtained by the proposed SAL-LSTM method and existing methods were evaluated in JJAS and shown in the Fig. 5. The Fig. 4 Shows that the proposed LSTM obtained higher performance compared to the other classifiers such as ANN, KNN, ELM in the rainfall prediction. The proposed SAL-LSTM method achieved higher performance due to its capacity to store the relevant features for the long time. The MAE of the proposed SAL-LSTM method is achieved as 1.82 and the existing LSTM method is achieved as 2.17. The analysis of the important features in the rainfall prediction helps the SAL-LSTM to achieve lower error value. The proposed SAL-LSTM is evaluated on JJAS data and compared with classifiers such as KNN, ANN, ELM and LSTM, as shown in Table 3.

The analysis shows that the proposed SAL-LSTM method has lower error value compared to other classifiers such as KNN, ANN, ELM and LSTM method. The proposed method has the advantage of storing the relevant features for long term and perform the classification.

The proposed SAL-LSTM method has 2.67 RMSE for JJAS data and LSTM method has 3.42 RMSE on rainfall predication.

The performance analysis of the various methods of OND data is compared in the Table 4. The Table 4 shows that the proposed LSTM method has the higher performance than the other prediction methods. The error metrics such as MAE, RMSE and MASE are evaluated for the three methods. The LSTM method

 DOI: 10.22266/ijies2022.0630.05

Received: December 1, 2021. Revised: January 25, 2022.

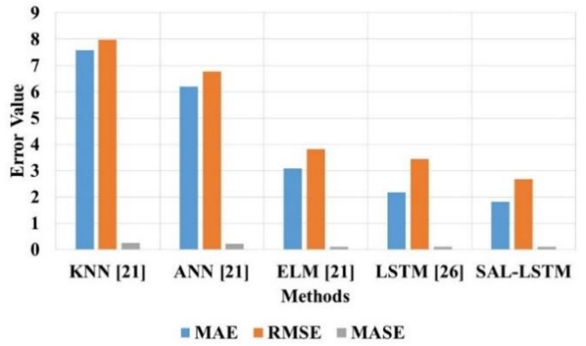

Figure. 5 The error value of several classifiers

Table 3. The performance analysis on JJAS data

Error Metrics	KNN [21]	ANN [21]	ELM [21]	LS TM	SAL-LSTM
MAE	7.58	6.18	3.07	2.17	1.82
RMSE	7.96	6.75	3.80	3.42	2.67
MASE	0.25	0.21	0.1	0.1	0.1

Table 4. The performance analysis in OND data

Error values	KNN [21]	ANN [21]	ELM [21]	LST M	SAL-LSTM
MAE	12.06	9.514	3.149	2.684	1.845
RMSE	17.07	11.54	6	5.237	3.986
MASE	0.24	0.19	0.05	0.017	0.015

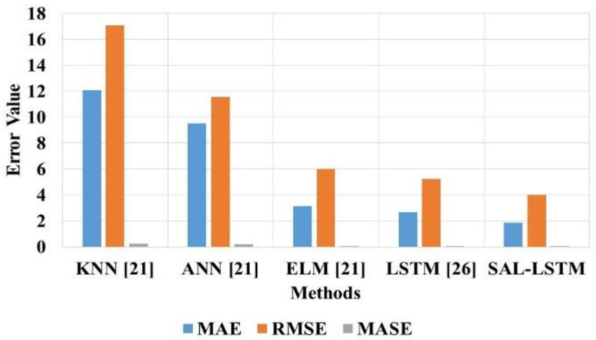

Figure. 6 The error measure of the existing and proposed method

has the highest performance due to the capacity of the method to handle the features for the long term. The RMSE of the SAL-LSTM method is achieved as 3.98 and the RMSE of the existing ANN method is achieved as 9.514.

The error value of the existing and proposed method is measure and comparison is shown as graphical representation in Fig. 6. The RMSE of the proposed SAL-LSTM method is achieved as 5.523 and the existing ELM method is achieved as 6. The proposed SAL-LSTM method has lower error value due to the method store the important features for the long term in the prediction. The LSTM, ANN and ELM have the considerable performance in the rainfall prediction.

Table 5. Performance analysis on Melbourne dataset

Architecture	RMSE	MAE	MASE
Dense Layer ANN [26]	1.11	0.2	0.19
GRU [26]	0.71	0.2	0.18
LSTM [26]	0.64	0.19	0.16
SAL-LSTM	0.52	0.12	0.02

Table 6. Error metrics analysis on Melbourne dataset

Methods	RMSE	MAE	MASE
KNN [21]	0.95	0.44	0.43
PCA – ANN [22]	0.93	0.36	0.35
CFBPNN [24]	0.85	0.37	0.32
PSO – SVM [25]	0.72	0.26	0.21
LSTM [26]	0.64	0.19	0.16
MHAL-LSTM [27]	0.63	0.17	0.15
AC-BiLSTM [28]	0.61	0.15	0.13
SAL-LSTM	0.52	0.12	0.02

4.3 Melbourne dataset

Melbourne dataset was used to evaluate the performance of the proposed SAL-LSTM method and compared with Dense layer ANN, Gated Recurrent Unit (GRU) and LSTM, as shown in Table 5.

The analysis shows that the SAL-LSTM method has the lower error value compared to Dense layer ANN, GRU and LSTM method. The existing LSTM model doesn't sufficiently handle the spatio-temporal data and this is important for the rainfall prediction model. The proposed SAL-LSTM method applies the attention layer to handle the spatio-temporal data to improve the efficiency of the rainfall prediction. The proposed SAL-LSTM method has 0.52 RMSE and existing LSTM has 0.64 RMSE.

The existing models such as KNN [21], PCA-ANN [22], CFBPNN [24], PSO-SVM [25], and LSTM [26] were evaluated in Melbourne dataset and compared with proposed SAL-LSTM model, as shown in Table 6. The table shows that proposed SAL-LSTM model has the lower MAE value compared to existing methods. The proposed SAL-LSTM model has the advantages of weighting the input features based on the softmax function that improves the efficiency of the model.

The proposed LSTM obtained low error value in two monsoon seasons compared with other existing ELM, ANN, KNN methods and the proposed LSTM achieved higher performance in 15 hidden node datasets.

5. Conclusion

Rainfall prediction is one of the challenging tasks due to the presence of the many environmental factors. In this research, LSTM method is used for the rainfall prediction based on Monsoon season. The Min-Max Normalization method is applied to normalize the missing data in the dataset. The LSTM is applied in the normalized data for monsoon and post-monsoon rainfall prediction in Kerala and its analysis the features in the data to predict the rainfall in the testing data. The proposed LSTM method is compared with the existing KNN, ANN and ELM methods in terms of MAE, RMSE, and MASE. The number of hidden layers is changed for the proposed and existing method to analyze its impact and the experimental results. The proposed LSTM method achieved higher MAE performance compared to the existing ELM, ANN as the proposed LSTM used 15 hidden layers for Rainfall prediction. The MAE of the proposed SAL-LSTM method has a lower value error of 1.724, whereas existing LSTM method showed the error value of 2.168. In the future, improvised SAL-LSTM can be used to solve vanishing gradient problem.

Conflicts of Interest

The authors declare no conflict of interest.

Author Contributions

The paper conceptualization, methodology, software, validation, formal analysis, investigation, resources, data curation, writing—original draft preparation, writing—review and editing, visualization, have been done by 1st author. The supervision and project administration, have been done by 2nd author.

References

[1] Y. Xiang, L. Gou, L. He, S. Xia, and W. Wang, "A SVR–ANN combined model based on ensemble EMD for rainfall prediction", *Applied Soft Computing*, Vol. 73, pp. 874-883, 2018.

[2] B. T. Pham, L. M. Le, T. T. Le, K. T. T. Bui, V. M. Le, H. B. Ly, and I. Prakash, "Development of advanced artificial intelligence models for daily rainfall prediction", *Atmospheric Research*, Vol. 237, p. 104845, 2020.

[3] J. D. Sierra and M. D. Jesus, "Long-term rainfall prediction using atmospheric synoptic patterns in semi-arid climates with statistical and machine learning methods", *Journal of Hydrology*, Vol. 586, p. 124789, 2020.

[4] T. Kashiwao, K. Nakayama, S. Ando, K. Ikeda, M. Lee, and A. Bahadori, "A neural network-based local rainfall prediction system using meteorological data on the Internet: A case study using data from the Japan Meteorological Agency", *Applied Soft Computing*, Vol. 56, pp. 317-330, 2017.

[5] K. J. Shou and J. F. Lin, "Evaluation of the extreme rainfall predictions and their impact on landslide susceptibility in a sub-catchment scale", *Engineering Geology*, Vol. 265, p. 105434, 2020.

[6] A. M. Bagirov, A. Mahmood, and A. Barton, "Prediction of monthly rainfall in Victoria, Australia: Clusterwise linear regression approach", *Atmospheric Research*, Vol. 188, pp. 20-29, 2017.

[7] D. N. Fernando, S. Chakraborty, R. Fu, and R. E. Mace, "A process-based statistical seasonal prediction of May–July rainfall anomalies over Texas and the Southern Great Plains of the United States", *Climate Services*, Vol. 16, p. 100133, 2019.

[8] K. S. Singh, S. Bonthu, R. Purvaja, R. S. Robin, B. A. M. Kannan, and R. Ramesh, "Prediction of heavy rainfall over Chennai Metropolitan City, Tamil Nadu", *Impact of Microphysical Parameterization Schemes Atmospheric Research*, Vol. 202, pp. 219-234, 2018.

[9] D. A. Ramu, S. A. Rao, P. A. Pillai, M. Pradhan, G. George, D. N. Rao, S. Mahapatra, D. S. Pai, and M. Rajeevan, "Prediction of seasonal summer monsoon rainfall over homogenous regions of India using dynamical prediction system", *Journal of Hydrology*, Vol. 546, pp. 103-112, 2017.

[10] K. C. Gouda, S. Nahak, and P. Goswami, "Evaluation of a GCM in seasonal forecasting of extreme rainfall events over continental India", *Weather and Climate Extremes*, Vol. 21, pp. 10-16, 2018.

[11] P. Arumugam and S. M. Karthik, "Stochastic Modelling in Yearly Rainfall at Tirunelveli District, Tamil Nadu", *Materials Today: Proceedings*, Vol. 5, No. 1, pp. 1852-1858, 2018.

[12] S. Cramer, M. Kampouridis, A. A. Freitas, and A. K. Alexandridis, "An extensive evaluation of seven machine learning methods for rainfall prediction in weather derivatives", *Expert Systems with Applications*, Vol. 85, pp. 169-181, 2017.

[13] Z. Wu, Y. Zhou, H. Wang, and Z. Jiang, "Depth prediction of urban flood under different rainfall return periods based on deep learning and data warehouse", *Science of The Total Environment*, Vol. 716, p. 137077, 2020.

[14] P. Singh and B. Borah, "Indian summer monsoon rainfall prediction using artificial neural network", *Stochastic Environmental*

Research and Risk Assessment, Vol. 27, No. 7, pp. 1585-1599, 2013.

[15] J. Abbot and J. Marohasy, "Input selection and optimisation for monthly rainfall forecasting in Queensland, Australia, using artificial neural networks", *Atmospheric Research*, Vol. 138, pp. 166-178, 2014.

[16] S. H. Pour, A. K. A. Wahab, and S. Shahid, "Physical-empirical models for prediction of seasonal rainfall extremes of Peninsular Malaysia", *Atmospheric Research*, Vol. 233, p. 104720, 2020, doi: 10.1016/j.atmosres.2019.104720.

[17] M. Ghamariadyan and M. A. Imteaz, "A Wavelet Artificial Neural Network method for medium-term rainfall prediction in Queensland (Australia) and the comparisons with conventional methods", *International Journal of Climatology*, Vol. 41, No. S1, pp. E1396-E1416, 2021, doi: 10.1002/joc.6775.

[18] A. S. Raj, D. H. Oliver, Y. Srinivas, and J. Viswanath, "Wavelet based analysis on rainfall and water table depth forecasting using Neural Networks in Kanyakumari district, Tamil Nadu, India", *Groundwater for Sustainable Development*, Vol. 5, pp. 178-186, 2017.

[19] S. Gao, Y. Huang, S. Zhang, J. Han, G. Wang, M. Zhang, and Q. Lin, "Short-term runoff prediction with GRU and LSTM networks without requiring time step optimization during sample generation", *Journal of Hydrology*, Vol. 589, p. 125188, 2020.

[20] Y. Dash, S. K. Mishra, S. Sahany, and B. K. Panigrahi, "Indian summer monsoon rainfall prediction: A comparison of iterative and non-iterative approaches", *Applied Soft Computing*, Vol. 70, pp. 1122-1134, 2018.

[21] Y. Dash, S. K. Mishra, and B. K. Panigrahi, "Rainfall prediction for the Kerala state of India using artificial intelligence approaches", *Computers & Electrical Engineering*, Vol. 70, pp. 66-73, 2018.

[22] Y. Dash, S. K. Mishra, and B. K. Panigrahi, "Predictability assessment of northeast monsoon rainfall in India using sea surface temperature anomaly through statistical and machine learning techniques", *Environmetrics*, Vol. 30, No. 4, p. e2533, 2019.

[23] H. Vathsala and S. G. Koolagudi, "Long-range prediction of Indian summer monsoon rainfall using data mining and statistical approaches", *Theoretical and Applied Climatology*, Vol. 130, No. 1-2, pp. 19-33, 2017.

[24] S. R. Devi, P. Arulmozhivarman, C. Venkatesh, and P. Agarwal, "Performance comparison of artificial neural network models for daily rainfall prediction", *International Journal of Automation and Computing*, Vol. 13, No. 5, pp. 417-427, 2016.

[25] B. T. Pham, L. M. Le, T. T. Le, K. T. T. Bui, V. M. Le, H. B. Ly, and I. Prakash, "Development of advanced artificial intelligence models for daily rainfall prediction", *Atmospheric Research*, p. 104845, 2020.

[26] J. Pudashine, A. Guyot, F. Petitjean, V. R. Pauwels, R. Uijlenhoet, A. Seed, M. Prakash, and J. P. Walker, "Deep Learning for an improved prediction of rainfall retrievals from commercial microwave links", *Water Resources Research*, Vol. 56, No. 7, p. e2019WR026255, 2020.

[27] J. Cao, Z. Li, and J. Li, "Financial time series forecasting model based on CEEMDAN and LSTM", *Physica A: Statistical Mechanics and its Applications*, Vol. 519, pp. 127-139, 2019.

[28] X. Zheng and W. Chen, "An Attention-based Bi-LSTM Method for Visual Object Classification via EEG", *Biomedical Signal Processing and Control*, Vol. 63, p. 102174, 2021.

[29] J. Zhao, D. Zeng, Y. Xiao, L. Che, and M. Wang, "User personality prediction based on topic preference and sentiment analysis using LSTM model", *Pattern Recognition Letters*, Vol. 138, pp. 397-402, 2020.