

Monthly Rainfall Forecasting Using BLSTM and GRU

Sathis Kumar K¹, Shobika S T²

¹ Assistant Professor Level III, Department of Computer Science and Engineering, Tamilnadu, India

² P.G and U.G Student, Department of Computer Science and Engineering, Tamilnadu, India

Abstract - Rainfall prediction is an important task due to the dependence of many people on it, especially in the agriculture sector. Prediction is difficult and even more complex due to the dynamic nature of rainfalls. We carry out monthly rainfall prediction over Coimbatore a region in the India. This examination contributes by giving a basic investigation and survey of most recent information mining procedures, utilized for precipitation expectation. The rainfall data were obtained from the National Center of Hydrology and Meteorology Department (NCHM) of India. We study the predictive capability with Linear Regression, Multi-Layer Perceptron (MLP), Convolution Neural Network (CNN), Long Short Term Memory (LSTM), Gated Recurrent Unit (GRU), and Bidirectional Long Short Term Memory (BLSTM) based on the parameters recorded by the automatic weather station in the region. Furthermore, proposed a BLSTM-GRU based model which outperforms the existing machine and deep learning models. From the six different existing models under study, LSTM recorded the best Mean Square Error (MSE) score of 0.0128. The proposed BLSTM-GRU model outperformed LSTM by 41.1% with a MSE score of 0.0075. Experimental results are encouraging and suggest that the proposed model can achieve lower MSE and high accuracy in rainfall prediction systems.

1. INTRODUCTION

Rainfall prediction has a widespread impact ranging from farmers in agriculture sectors to tourists planning their vacation. Moreover, the accurate prediction of rainfall can be used in early warning systems for floods and an effective tool for water resource management. Despite being of paramount use, the prediction of rainfall or any climatic conditions is extremely complex. Rainfall depends on various dependent parameters like humidity, wind speed, temperate, etc., which

vary from one geographic location to another; hence, one model developed for a location may not fit for another region as effectively. Generally, rainfall can be predicted using two approaches. The first is by studying all the physical processes of rainfall and modeling it to mimic a climatic condition. However, the problem with this approach is that the rainfall depends on numerous complex atmospheric processes which vary both in space and time. The second approach is using pattern recognition. These algorithms are decision tree, k-nearest neighbor, and rule-based methods.

2. PROJECT DESCRIPTION

EXISTING SYSTEM

In our existing framework we will utilize CNN calculation for the expectation of precipitation. Likewise, we assess the presentation aftereffect of calculation dependent on mean outright blunder, mean squared mistake and root mean squared error. It upgrades the exhibition of the general characterization results. The data is survived from the weather station database and data's are selected according to corresponding location and the non-usable data's are removed in missing data removal phase and the categorical data's area encoded using convolutional neural network algorithm in order to convert them into integer data type which makes easy to predict rainfall accuracy in which both this operation in done in data pre - processing phases. The data retrieved from data pre - processing is trained to predict rainfall in corresponding location in which 70% of data's are taken to training and other 30% of data's are taken for testing.

PROPOSED SYSTEM

In our proposed framework we will utilize BLSTM and GRU calculation for the expectation of precipitation. This framework brings about high execution and provides precise expectation results and furthermore reduces the data loss and the predisposition of the surmising because of the various evaluations.

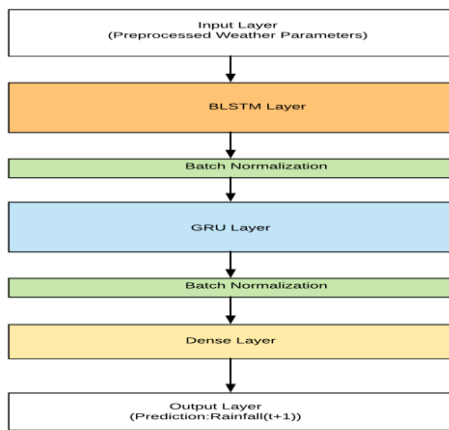


Fig 1 Proposed Systems

The proposed model is composed of 7 layers including the input and output layers. The embedding is generated by the Bidirectional Long Short Term Memory (BLSTM) and Gated Recurrent Unit (GRU) layer. The batch normalization is used for normalizing the data, and the dense layer performs the prediction.

3. MODULES AND DIAGRAMS

Modules applied in rainfall forecasting for providing an efficient result are

- Data Selection and Loading
- Data Preprocessing
- Evaluation Metrics
- BLSTM
- GRU
- BLSTM and GRU

DATA SELECTION AND LOADING

The data selection is the process of selecting the data predicting the rainfall. In this project, the rainfall dataset is used for predicting rainfall. The dataset contains the information about the monthly wise data of rainfall with corresponding location.

DATA PREPROCESSING

The rawdata originally contained eight parameters, but some of the parameters contained a lot of missing and noisy values. The weather parameters that contained a lot of empty records were dropped from the dataset. The dataset also had different random representations for the null value, which was standardized during preprocessing. The missing values in the selected parameters were resolved by taking the mean of all the values occurring for that particular day and month. Outliers are records that significantly differ from other observed values. The outliers were detected using a box-and-whisker plot as well as the k-means clustering algorithm and were resolved using the mean technique. After preprocessing, the data are reshaped into a tensor format.

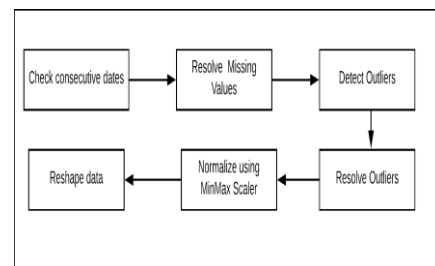


Fig 2 Data Preprocessing

EVALUATION METRICS

The study used both qualitative and quantitative metrics to calculate the performance of different models. The formulae for RMSE, MSE, Pearson Correlation Coefficient, and R^2 were used as a scoring function.

Name	Formula
MSE	$\frac{1}{n} \sum_{t=1}^n (x_t - y_t)^2$
RMSE	$\sqrt{\frac{1}{n} \sum_{t=1}^n (x_t - y_t)^2}$
R^2	$1 - \frac{\sum_{i=0}^{n_{samples}-1} (x_i - y_i)^2}{\sum_{i=0}^{n_{samples}-1} (x_i - \bar{x})^2 + \sum_{i=0}^{n_{samples}-1} (y_i - \bar{y})^2}$
Correlation	$\frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$

From the above, x_t is the model simulated monthly rainfall, y_t is the observed monthly, \bar{x} and \bar{y} are their arithmetic mean, and n is the number of data points.

BLSTM

LSTM is the most popular model in time series analysis, and there are many variants such as unidirectional LSTM and BLSTM. For our study, the Many-to-One (multiple input and one output) variation of LSTM was used to take the last 12 months' weather parameters and predict the rainfall for the next month/ Unidirectional LSTM process data are based on only past information. Bidirectional LSTM utilizes the most out of the data by going through time-steps in both forward and backward directions. It duplicates the first recurrent network in the architecture to get two layers side by side. It passes the input, as it is to the first layer and provides a reversed copy to the second layer. Although it was traditionally developed for speech recognition, its use has been extended to achieve better performance from LSTM in multiple domains. An architecture consisting of two hidden layers with 64 neurons in the first layer and 32 neurons in the second layer recorded the best result on the test dataset, with MSE value of 0.01, a coefficient value of 0.87, and R^2 value of 0.75.



Fig 3 Many to One LSTM

GRU

The Gated Recurrent Unit was developed by Cho et al. in 2014. GRU performances on certain tasks of natural language processing, speech signal modelling, and music modelling are similar to the LSTM model. The GRU model has fewer gates compared to LSTM and has been found to outperform LSTM when dealing with smaller datasets. To solve the vanishing gradient problem of a standard RNN, GRU consists of an update and reset gate, but unlike the LSTM it lacks a dedicated output gate. The update gate decides how much of the previous memory to keep, and the reset gate determines how to combine the previous memory with the new input. Due to fewer gates, they are computationally less demanding compared to LSTM and are ideal when there are limited computational resources. GRU with two hidden layers consisting of 12 neurons in the first layer and 6 neurons in the second outperformed other architectures, with an MSE score of 0.02, a correlation value of 0.83, and R^2 value of 0.66.

BLSTM AND GRU

In this model, pre-processed weather parameters are fed into the BLSTM layer with 14 neurons. This layer reads data in both forward and backward directions and creates an appropriate embedding. Batch normalization is performed on the output of the BLSTM layer to normalize the hidden embedding before passing it to the next GRU layer. The GRU layer contains half the number of neurons as the BLSTM layer. The GRU layer has fewer cells and is able to generalize the embedding with relatively lower computation cost. The data are again batch normalized before sending to the final dense layer. The final layer has just one neuron with a linear activation function, and it outputs the predicted value of monthly rainfall for $T + 1$ (next month), where T is the current month.

For our study, the Many-to-One (multiple input and one output) variation of LSTM was used to take the last 12 months' weather parameters and predict the rainfall for the next month. The activation function used in both BLSTM and GRU is the default tanh function, and the optimizer used was Adam. The architecture was fixed after thoroughly hyper-tuning the parameters. Hyperparameter tuning was performed through a randomized grid search and heuristic knowledge of the programmer.

3. EXPERIMENTAL RESULT

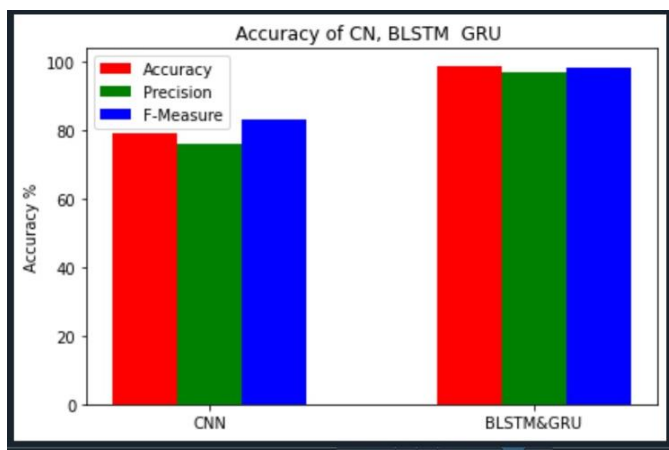


Fig 4 Final Result

4. CONCLUSIONS

The study of deep learning methods for rainfall prediction is presented in this paper, and a BLSTM-GRU based model is proposed for rainfall prediction. Finally the combination of BLSTM and GRU layers performed much better than all the other models under study for this dataset. Its MSE score of 0.007 was 41.1% better than LSTM. Furthermore, the proposed model presented an improved correlation value of 0.93 and R2 score of 0.87.

FUTURE WORK

We aim to improve the performance of our prediction model by incorporating patterns of global and regional weather such as sea surface temperature, global wind circulation, etc. We also intend to explore the predictive use of climate indices and study the effects of climate change on rainfall patterns.

REFERENCES

- Toth, E.; Brath, A.; Montanari, A. Comparison of short-term rainfall prediction models for real-time flood forecasting. *J. Hydrol.* 2000, 239, 132–147.
- Jia, Y.; Zhao, H.; Niu, C.; Jiang, Y.; Gan, H.; Xing, Z.; Zhao, X.; Zhao, Z. A webgis-based system for rainfall-runoff prediction and real-time water resources assessment for beijing. *Comput. Geosci.* 2009, 35, 1517–1528.
- Walcott, S.M. *Thimphu.* Cities 2009, 26, 158–170.
- Abhishek, K.; Kumar, A.; Ranjan, R.; Kumar, S. A rainfall prediction model using artificial neural network. In *Proceedings of the Control and System Graduate Research Colloquium (ICSGRC), Shah Alam, Malaysia, 16–17 July 2012;* pp. 82–87.
- Darji, M.P.; Dabhi, V.K.; Prajapati, H.B. Rainfall forecasting using neural network: A survey. In *Proceedings of the 2015 International Conference on Advances in Computer Engineering and Applications (ICACEA), Ghaziabad, India, 19–20 March 2015;* pp. 706–713.
- Kim, J.-H.; Kim, B.; Roy, P.P.; Jeong, D.-M. Efficient facial expression recognition algorithm based on hierarchical deep neural network structure. *IEEE Access* 2019, 7, 41273–41285.
- Mukherjee, S.; Saini, R.; Kumar, P.; Roy, P.P.; Dogra, D.P.; Kim, B.G. Fight detection in hockey videos using deep network. *J. Multimed. Inf. Syst.* 2017, 4, 225.