International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 05 | May - 2023 SJIF 2023: 8.176 ISSN: 2582-3930

LSTM USING RAINFALL PREDICTION BASED ON AGRICULTURE

Pavan Kalyan Ch
Dept. of Computer Science &Engineering
Prathyusha Engineering College
Tiruvallur, Tamilnadu
soliderpavankalyan14@gmail.com

Venkata Sai Charan M
Dept. of Computer Science &Engineering
Prathyusha Engineering College
Tiruvallur, Tamilnadu
charanmenta169@gmail.com

Ms. Shimona E
Assistant professor, Dept. of Computer Science &Engineering
Prathyusha Engineering College
Tiruvallur, Tamilnadu
shimona.cse@prathyusha.edu.in

Charan Teja C
Dept. of Computer Science &Engineering
Prathyusha Engineering College
Tiruvallur, Tamilnadu
charantejareddy21@gmail.com

Vishnu Vardhan D
Dept. of Computer Science &Engineering
Prathyusha Engineering College
Tiruvallur, Tamilnadu
vishnudevineni339@gmail.com

***Abstract*— The most difficult task of meteorology is to predict rainfall. In our study, we proposed an amount of rainfall prediction model that can be easily determined using machine learning and LSTM techniques. This is an advanced method to find out the rainfall. The deep learning approach is most valuable for this type of method implementation and its accuracy finds out. A long short-term memory algorithm is applied to memory sequence data measurement and calculate previous data very fast and create the best prediction. The people of this country are mostly dependent on agriculture so that this prediction system is very necessary. Timely rainfall assessment will increase crop yields and reduce costs in agriculture. Considering all these factors, we have created our model which will help us to determine the amount of rainfall. We have collected data from 6 regions to do this. To predict, wehave taken 6 parameters (temperature, dew point, humidity, wind pressure, wind speed, and wind direction). After analyzingall our data, we got 76% accuracy in our work. We also focus ona vast dataset in long time weather for the better result.**

Keywords*— *Rainfall Forecast, LSTM, RNN, Machine learning

I. INTRODUCTION

Rainfall is a very large part of the climate system which plays a necessary role in the formation of nature. In some parts of Bangladesh such as Jessore, Kushtia, Rajshahi, Pabna, Dinajpur, Bogra, Natore, Jhenaidah, Chadpur rainfall is very important for agricultural work. But many times, more rain and unwanted heavy rains bring a lot of misery for us. Excessive rainfall often leads to floods which often cause a lot of damage and death. But if we get an accurate forecast of rainfall, we can reduce the amount of damage. But the forecast of rain can bring a good direction for us, if not, it is very accurate. With the utmost importance to this accuracy, we have made a decision based on our research which will bring many benefits in the future. After analyzing all the factors, we have applied the most popular method of machine learningto get perfect accuracy. Here we have tried to develop a special model and arrange some important data analysis. In our study, we mainly focus on Long Short Term Memory and recurrent agriculture (RNN). We used real time frequency, deep learning latest algorithm, and probability theorem for prediction accuracy. Short time memory is a major part of this study. We are also developing a model of prediction table where shows the accuracy of the forecast. Time series forecast analysis is the latest approach in deep learning and we are focus to develop our agriculture and flood effect damage remove the system.

II. RELATED WORK

Rainfall is one of nature's most remarkable, vital, and essential factors that is essential for the urban economy. Floods created for rainfall, which destroys roads and crops, greatly reduce the quality of life. Many approaches have been proposed for accurately determining the forecast that would turn an urban area into a beautiful economic area [1]. There are many states in India where people are dependent on rainfall for agriculture, and there are also states where land has turned into a desert due to lack of rainfall. Floods and landslides caused by excessive rainfall in many states also cause many problems, reducing the standard of living of these estates and adversely affecting the economy. Rainfall does not adhere to any specific time, so it is very challenging to determine the right time. There are many categories of artificial agricultures and they are predicted by maintaining time-series in a completely uncommon way through rainfall analysis [2]. Water is one of the natural resources. Precipitation data is very important for an area. If climate models are created mathematically, then much progress can be made in rainfall prediction. Rainfall models of different regions can be created by establishing relationships through large-scale weather and local observations [3]. Floods are one of the reasons for the loss ofhuman life and economic losses which puts our social life at risk and disrupts our daily life. To get rid of this, we do not have to face any loss if we observe the flood of the river in advance and present its correct mathematical picture but this process is very complicated because it depends on the climate, the direction of river flow, rainfall, soil, location. This processis much easier based on hydrology. But still, the error remains in the conventional model [4]. Tourism flow is a process that is difficult to diagnose because it is not linearly affected. Many approaches LSTM NN, ARTMA, BPNN have been used in various steps to determine the correct tourism flow [10]. Artificial agricultures are much more efficient than any other method for diagnosing streamflow and it is much more

developed than ever before. If the WGRNN and FFNN models are compared, then the WGRNN model is much more effective because the calculation is easier with less time [11].

III. Proposed Model

In our research, we have proposed a sample model of the complete working process. Whole working process Fig. 1, we are doing separate it 4 sections. In the first section, it is an input section part. Here we collect many graphs with 4 major components like Wind direction, temperature, wind speed, and pressure. We collect many sets for LSTM processing. Sample 1, sample 2, Samples Nth of data set.

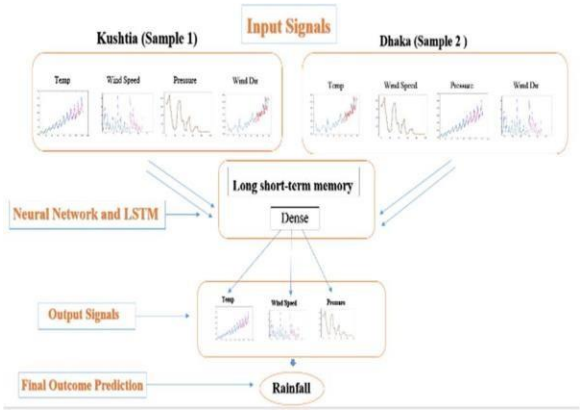

Fig. 1. A full simulation process of rainfall prediction.

The middle portion indicates the LSTM and RNN Processing section. All sample data set processed by LSTM coding programming and analysis. At the end of the part, our main target is to find out rainfall prediction.

IV. Method and Data Analysis

To complete the whole study, we have prepared a summary methodology for all the work. Here algorithm analysis, mapping model, the circuit has been built up. With this, we have presented a visual of data processing and analysis in this section.

A. Long Short-Term Memory

Long Short Term Memory which is major and latest algorithm of deep learning. In our study, we mainly focus on the memory block [5]. Because of, for data analysis and its accurate prediction recurrent sort memory is the main part of rainfall time calculation. It contains 3 fundamental part in gates: (1) Input Gate (2) Output Gate (3) Forget Gate. To identify the forecast prediction, we use this gate with the info flow. Gates in LSTM are the sigmoid activation functions i.e. they output a worth between 0 or 1. "0" means the gates are blocking everything. "1" means gates are allowing everything to undergo it. Basic timestamp LSTM Memory circuit model is shown in Fig. 2, for rainfall prediction.

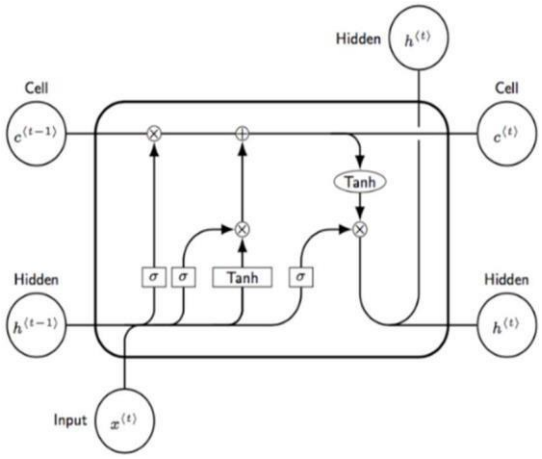

Fig. 2. A basic timestamp {t} LSTM Memory circuit model for rainfall

The main 3 equations of LSTM: The first equation which is Input (i) Gate, tells us what new information we're going to store in the cell state.

$$\mathrm{i} = \sigma(\mathrm{W}_{xi}x_t + \mathrm{W}_{hi}h_{t-1} + \mathrm{W}_{ci}C_{t-1} + b_{\mathrm{i}}) \quad (1)$$

We can find out through Hidden Gate (2) what information we should keep.

$$\mathbf{f} = \sigma(\mathrm{W}_{xf}x_t + \mathrm{W}_{hf}h_{t-1} + \mathrm{W}_{cf}C_{t-1} + b_{\mathbf{f}}) \quad (2)$$

Lastly, for using the output gate which is used to provide the activation to the final output (o) of the LSTM block at timestamp 't'.

$$\mathrm{o}_t = (\mathrm{W}_{xo}\mathrm{X}_t + \mathrm{W}_{ho}h_{t-1} + \mathrm{W}_{co}C_t + b_{\mathrm{o}}) \quad (3)$$

To get the memory vector for the current timestamp (c{t}) the temperature is calculated. To absorb benefit from both methods, we are using the neural layers to extract the local features of the raw input and then use them to the LSTM networks [8]. In our study, predictors from the past 15 to 25 days are used to count daily rainfall.

B. Recurrent Agricultures and Tanh Activation

RNN processes predict the rainfall using the wind speed, pressure, the temperature of vectors one by one in sequence. While processing this variable, it passes the previously hidden data status to the next step in the sequence. The hidden state serves as the memory of the agriculture. It contains all the information related to the old data that the network has seen before. The basic concept of LSTMs is the cell state and its various gates. The cell state acts as a transport highway that makes relative transfers.

$$h_t = O_t\tanh(C_t) \quad (4)$$

Gates has sigmoid activation. Sigmoid activation is similar to Tanh activation shown in Fig. 3. The value of -1 and 1 squishes between 0 and 1 instead of squish [9]

When vectors flow through a agriculture, it goes throughmany metaphors due to various mathematical activities that help to find the accurate accuracy and predictivedata tables.

C. *Model Architecture*

Time series forecasting with LSTM is a modern approach to build a rapid model of forecasting. To the best study, we make the architecture model after time series analysis. A great surprise as agricultures are known to be ready to learn complex nonlinear relationships and therefore the LSTM is probably the foremost successful sort of recurrent agriculture that's capable of directly supporting multivariate sequence prediction problems.

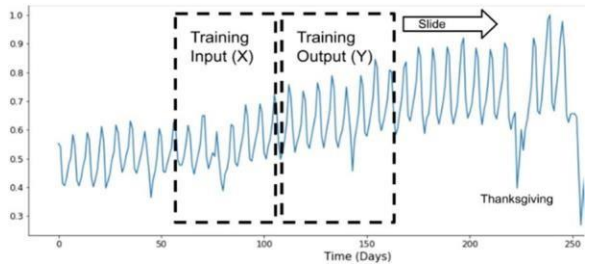

Fig. 4. Time Frequency Architecture for Rainfall

A training datasets is created. Dataset splitting process is done by using datamining technique [6, 7]. Splitting the dataset by district based rainfall historical data into sliding windows of input (X) and output (Y) variables which is shown in Fig 4. The particular size of the look-back and forecast view used in the experiments is not specified in the research model.Time series data is scaled by simplify and monitoring per batch of samples and every input series is de-trended, but notdepersonalized.

D.Data Collection

We obtained weather data set from Bangladesh Meteorological Department (BMD) [12]. In the raw weather dataset, six measured parameters are temperature, dew-point, humidity, wind pressure, wind speed and wind direction. Last 1 month of meteorological data during monsoon season (1 august to 31 august) for year 2020 were collected and used for training. In this analysis, data variations have been used to test the robustness and reliability of proven modeling for extreme events. The attribute values have taken in numerical however,our model requires computational categorical values. As per requirement, the value set has been separated and values of attributes have been transformed into categorical data. Table1, illustrates the attributes that we have evaluated and preprocessed.

TABLE I. TABLE TYPE STYLES

Attribute	Type	Description
Temperature	Numerical	°C(Degree Celsius)
Dew-point	Numerical	°C(Degree Celsius)
Humidity	Numerical	%(Percentage)
Wind Pressure	Numerical	In Hectopascal
Wind Speed	Numerical	In Kilometer per Hour
Wind-Direction	Numerical	In Degree

V. RESULT AND DISCUSSION

After analyzing all our data using LSTM and RNN, we found 76% accuracy in our work. We have data of 12 regions, from where we make these predictions. Fig. 5, shows the amount of rainfall in Dhaka city from 2000 to 2014. From this figure, we can figure out the actual rainfall and forecast rainfall.

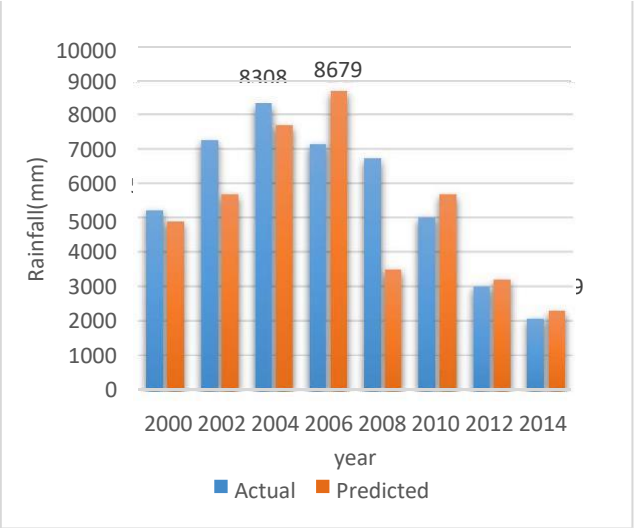

Fig. 5. Yearly prediction of rainfall quantity

Using this model, we also predict the monthly rainfall quantity, which is shown in Fig. 6.

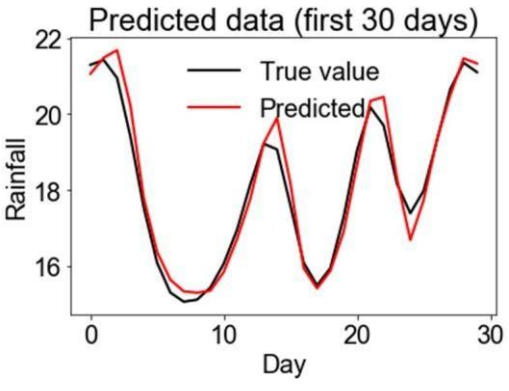

Fig. 6. Monthly rainfall quantity predictio

VI. Conclusion

In this study, our model based on LSTM and RNN has been developed to predict the amount of rainfall on a monthly basis and on an annual basis. For accurate forecasting, we use wind direction, temperature, wind speed, and pressure data to predict the amount of rainfall in a year or a month. Rainfall is one of the most essential parts of agriculture. Farmers are mostly dependent on rain, although there are many modern technologies [13]. But they are expensive and difficult to manage. So farmers will benefit if they know the weather forecast early and know when or how much rainfall will occur. In this work, our main goal is to find solutions to how farmers can benefit from using rainfall forecasting.

Acknowledgment

To complete our research work Daffodil International University help us. We use NLP and Machine learning lab for better result.

References

[1] Le V.M., Pham B.T., Le TT., Ly HB., Le L.M. (2020) Daily Rainfall Prediction Using Nonlinear Autoregressive Agriculture. In: Sharma D., Balas V., Son L., Sharma R., Cengiz K. (eds) Micro- Electronics and Telecommunication Engineering. Lecture Notes in Networks and Systems, vol 106. Springer, Singapore. https://doi.org/10.1007/978-981-15-2329-8_22

[2] Poornima, S.; Pushpalatha, M. Prediction of Rainfall Using IntensifiedLSTM Based Recurrent Agriculture with Weighted Linear Units. Atmosphere 2019, 10, 668.

[3] Miao, Q.; Pan, B.; Wang, H.; Hsu, K.; Sorooshian, S. Improving Monsoon Precipitation Prediction UsingCombined Convolutional and Long Short Term Memory Agriculture. Water 2019, 11, 977.

[4] Le, X.-H.; Ho, H.V.; Lee, G.; Jung, S. Application of Long Short-Term Memory (LSTM) Agriculture for Flood Forecasting. Water 2019,11, 1387.

[5] W. Kong, Z. Y. Dong, Y. Jia, D. J. Hill, Y. Xu and Y. Zhang, "Short-Term Residential Load Forecasting Based on LSTM Recurrent Agriculture," in IEEE Transactions on Smart Grid, vol. 10, no. 1, pp. 841-851, Jan. 2019, doi: 10.1109/TSG.2017.2753802

[6] I. M. Talha, I. Salehin, S. C. Debnath, M. Saifuzzaman, N. N. Moon and F. N. Nur, "Human Behaviour Impact to Use of Smartphones with the Python Implementation Using Naive Bayesian," 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Kharagpur, India, 2020, pp. 1-6, doi: 10.1109/ICCCNT49239.2020.9225620.

[7] I. Salehin, I. M. Talha, N. N. Moon, M. Saifuzzaman, F. N. Nur and M. Akter, "Predicting the Depression Level of Excessive Use of Mobile Phone: Decision Tree and Linear Regression Algorithm," 2020 International Conference on Sustainable Engineering and Creative Computing (ICSECC), 16-17 December 2020, President University, Indonesia., in press.

[8] R. Jiao, T. Zhang, Y. Jiang and H. He, "Short-Term NonResidential Load Forecasting Based on Multiple Sequences LSTM Recurrent Agriculture," in IEEE Access, vol. 6, pp. 59438-59448, 2018, doi: 10.1109/ACCESS.2018.2873712.

[9] Godin, F., Degrave, J., Dambre, J., & De Neve, W. (2018). Dual Rectified Linear Units (DReLUs): A Replacement for Tanh Activation Functions in QuasiRecurrent Agricultures. Pattern Recognition Letters. doi:10.1016/j.patrec.2018.09.006.

[10] Li, Y., & Cao, H. (2018). Prediction for Tourism Flow based on LSTM Agriculture. Procedia Computer Science, 129, 277– 283. doi:10.1016/j.procs.2018.03.076

[11] Kişi, Ö. (2011). A combined generalized regression agriculture wavelet model for monthly streamflow prediction. KSCE Journal of Civil Engineering, 15(8), 1469–1479. doi:10.1007/s12205-011-1004-4