PREDICTION OF METEOROLOGICAL DROUGHT OVER WESTERN RAJASTHAN USING DEEP-LEARNING TECHNIQUES

A MINI-PROJECT REPORT

submitted by

JAYESH RAJ, TKM20MEAI09

to

the APJ Abdul Kalam Technological University

in partial fulfillment of the requirements for the award of the Degree

of

Master of Technology

In

Mechanical Engineering (Artificial Intelligence)

CENTRE FOR ARTIFICIAL INTELLIGENCE

T.K.M. COLLEGE OF ENGINEERING, KOLLAM

JULY, 2021

CENTRE FOR ARTIFICIAL INTELLIGENCE
T.K.M COLLEGE OF ENGINEERING, KOLLAM

CERTIFICATE

This is to certify that the report entitled **'PREDICTION OF METEOROLOGICAL DROUGHT OVER WESTERN RAJASTHAN USING DEEP LEARNING TECHNIQUES'** submitted by **Jayesh Raj** to the APJ Abdul Kalam Technological University in partial fulfillment of the requirements for the award of the Degree of Master of Technology in Mechanical Engineering (Artificial Intelligence) is a bonafide record of the project work carried out by her under my guidance and supervision. This report in any form has not been submitted to any other University or Institute for any purpose.

Dr.Adarsh S
Project guide
Associate professor
Deptartment of Civil Engineering
TKMCE

Prof. Sumod Sundar
Project Coordinator
Assistant Professor
Centre of Artificial Intelligence,
TKMCE

Dr. Imthias Ahamed T P
Head of Department
Professor
Centre of Artificial Intelligence,
TKMCE

ACKNOWLEDGEMENT

Ithank **God Almighty** who strengthen us to complete the project. Iexpress our sincere gratitude to Principal **Dr.Shahul Hameed T A,** for providing the lab and necessary facilities. Iexpress our gratitude to **Dr. Imthias Ahamed T P**, Head of Centre for Artificial Intelligence, TKM College of Engineering, Kollam, for all necessary help extended for us in the fulfillment of this work. Iare immensely indebted to our guides **Dr.Adarsh S** for their constructive guidance and advices.

I show our extreme gratitude to all Faculties and Technical staffs in Centre for Artificial Intelligence, for providing all the help and necessary facilities to present this thesis.

Our deep hearted cheers to our parents and all our friends who extended their support and co-operation towards the successful presentation of the thesis.

Place: Kollam
Date:

Jayesh Raj, TKM20MEAI09

ABSTRACT

Over years, natural calamities like drought have taken a huge toll on human life and resources. As the prediction methods increase, the effects of natural calamities can be reduced to an extent by preplanning and providing warnings to the people. Metrological drought indices like standardized precipitation index are used to identify drought and its severity level. By forecasting these indices, the occurrences of drought are predicted using the prediction models which help the society to take preventive measures due to the effect of drought. Many research works on prediction majorly focused on statistical methods such as Holt–Winters and ARIMA, but these methods lack accuracy to provide long-term forecasts. However, with advances in the area of machine learning especially artificial neural networks and deep neural networks, there seems to be a method to predict drought in the long term with a good accuracy. Long short-term memory is used in recurrent neural network to predict the drought indices which handle the real-time nonlinear data well and good that can help authorities better prepare and mitigate natural disasters. In this paper, I take 3-month prediction for LSTM.

Contents

List of Figures

Abbreviations

1. LSTM - Long Short Term Memory
2. SPI- Standardized Precipitation Index
3. PDSI - Palmer Drought Severity Index
4. MSE - Mean Square Error
5. RMSE - Root Mean Square Error
6. MAE - Mean Absolute Error

Chapter 1

Introduction

Precipitation is a field that is random in character and as such makes drought prediction a complex task. Drought materializes from a deficiency of rainfall over a period of time. Drought over short timescales (months) characterize meteorological drought, whereas long-term scales (years) showcase hydrological drought. Drought causes immense damage to the environment, economy and society. There is an abrupt increase in fires and deflation intensity loss of biodiversity and introduction of pests and diseases.

Agricultural losses result in higher suicide rates among farmers, higher cost of food production, lower hydrological energy output and depleted water supply and tourism. They are also responsible for excessive heat waves, limitation of water supplies for consumption, high stress due to failed harvests, etc. Therefore, there is a vital requirement to give accurate prediction of drought occurrence especially for a longer timescale.

1.1 Overview of SPI

The standardized precipitation index (SPI) was proposed in order to help monitor relative wetness and dryness over multiple timescales.Short timescales imply that the SPI is closely related to soil moisture, while at longer timescales it indicates groundwater and reservoir storage. SPI can be used across regions with differing cli- matic conditions. This results from the quantification of observed precipitation as a selected probability distribution function that is modeled over raw precipitation data. SPI can be used over 1–36-month timescale and can be interpreted as the number of standard deviations by which the observed anomaly deviates from long-term mean.Several indices pertaining to drought were formulated and used all over the scientific world. These were based on percentage of rainfall and percentile values and sometimes were very complex like the PDSI. The main aim of developing this index was to get an index that was easy in complexity and calculation so that they were applicable in regions all over the world.The deficit of rainfall has varied effects on various water-related resources and phenomenon such as stream flow, groundwater and soil moisture. This resulted in them formulating the SPI. The index is powerful, flexible and fairly easy to use and is simple to calculate. Precipitation of the time series is the only time series required.Its effectiveness is broad with respect to the phe- nomenon of rainfall as it can analyze drought and floods in equal measures. The SPI formulation was based on the purpose to implement the precipitation values for several timescales. From the timescales, Ican obtain the effects of drought on whether certain water bodies and resources are available or viable for use. The SPI calculation for the place is calculated on the long-term precipitation values foR the period one desires. This long-term record is fit into a probability distribution, which is then transformed into a normal distribution. This implies that the mean SPI for the location and period desired is zero. Positive SPI values show greater than median precipi- tation, while negative values show less than median rain- fall. Because the SPI is normalized over a standard distribution, wetter and drier climates can be represented in the same way. Wet periods, whether it is slight or heavy rainfall, can also be monitored using the SPI.Drought classification is based on Figure 1.1 . A drought is set to have occurred if the SPI value is continuously negative and reaches a magni-tude of -1 or less. The event is supposed to end when the SPI reaches a positive value. SPI has flexibility and can be calculated for several time periods or timescales. Shorter time period or scale SPI is known to provide early warning of drought and also help predict very accurately the drought intensity. It provides a means of comparison of locations in different climates. Since it is probabilistic in nature, it acquires historical basis that would help in making important decisions. However, it is based on explicitly only the precipitation parameter and has no soil–water balance indicator, and no associated ratios of evapotranspiration/potential evapotranspiration could be evaluated.

2.0+	Extremely wet
1.5 to 1.99	Very wet
1.0 to 1.49	Moderately wet
− .99 to .99	Near normal
− 1.0 to − 1.49	Moderately dry
− 1.5 to − 1.99	Severely dry
− 2 and less	Extremely dry

Figure 1.1: SPI range for drought

In this research paper I Selected drought prone sub-divisions of Western Rajasthan and collected the monthly rainfall data from 1950 to 2016.Long short-term memory is used in recurrent neural network to predict the drought indices which handle the real-time nonlinear data well and good that can help authorities better prepare and mitigate natural disasters. I take 3-month prediction of LSTM.

Chapter 2

Methodology

2.1 Proposed System

In my work I propose a “LSTM for predicting the drought over Western-Rajasthan”. The model assists the R software to convert the precipitation values to SPI values at an early stage. SPI package is the mathematical tool used to extract the SPI values(time scale-3)

Steps undertaken:

- Data set is collected from IITM Pune-36 Meteorological(1958-2016)
- Arranged the precipitation data according to monthly
- extract the SPI value using SPI package in R
- Converted to csv file
- Dataset is splited into 75 percentage for training and remaining
- Then this csv file is loaded to LSTM
- And then SPI value is prediction

Then,the obtained SPI value is checked with SPI range for drought ,Figure 1.1.Hence the drought prediction is done.

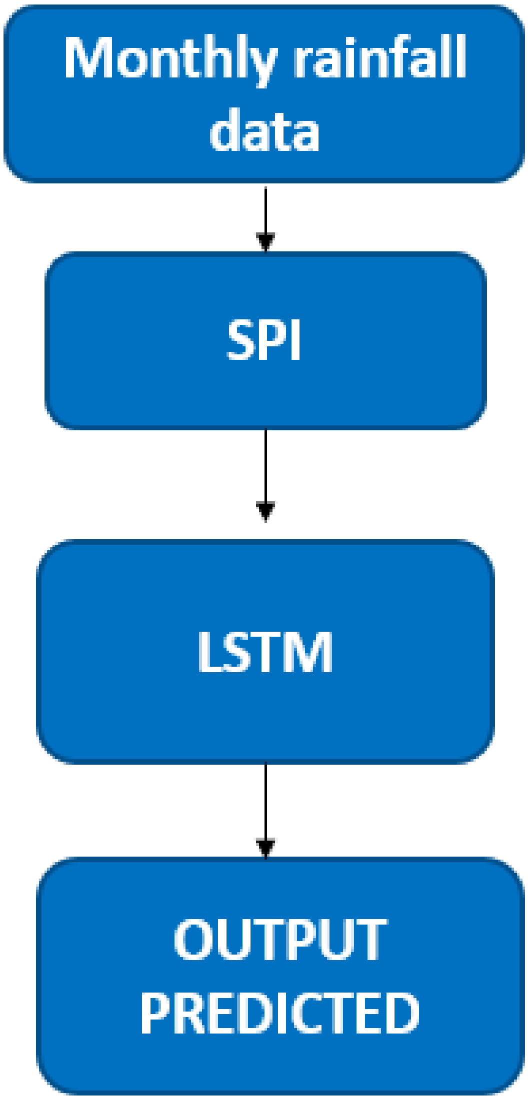

Figure 2.1: System flow chart

2.2 LSTM

Long short-term memory (LSTM) is a recurrent neural network (RNN) architecture (an artificial neural network) proposed. RNNs can capture the dynamics of sequences via cycles in the network. But some RNNs suffer from the vanishing and exploding gradients problem in which gradients are either squashed to zero or increase without bound during backpropagation through a large number of time steps. LSTM is introduced primarily to overcome the problem of vanishing gradients. LSTM network is well suited to learn from experience to classify process and predict time series when there are time lags of unknown size and bound between important events. LSTM is able to model long-term dependencies by using a memory unit called cell state. It has a chain like structure, having four gates which are implemented using the logistic function. All the four gates take the previous state as input along with the current input. The role of each gate is as follows: Forget gate controls the extent to which the value remains in memory, input gate allows the flow of new values in the memory, candidate gate generates the new update for the cell state, and finally, output gate allows the value in memory to compute the output activation of the block further. Forget gate, input gate and output gate use sigmoid function to perform its task, whereas the candidate gate and output gate use tanh function. All the logistic functions are computed by applying a weight and a bias to trigger the neurons and normalize the inputs. Every neuron in the hidden layer of recurrent neural network is implemented with LSTM unit and undergoes for number of states for prediction. The above-mentioned functionalities of LSTM are shown in Fig. 2.2.

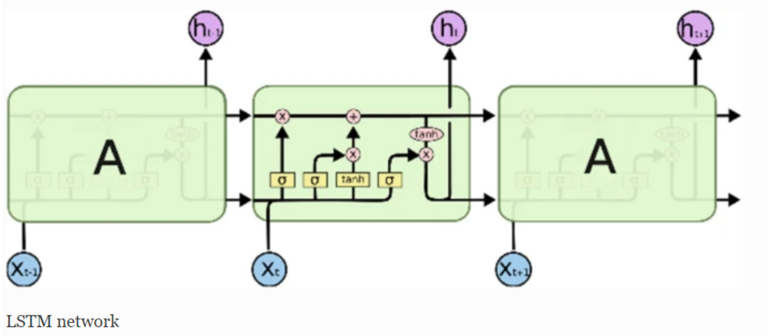

Figure 2.2: LSTM

The information given depicts the forward pass and backward pass in LSTMs. In terms of the forward pass, the LSTM can learn when to let activation into the internal state. As long as the input gate takes value zero, no activation can get in. Similarly, the output gate learns when to let the value out. When both gates are closed, the activation is trapped in the memory cell, neither growing nor shrinking, or affecting the output at intermediate time steps. In terms of the backwards pass, the constant error carousel enables the gradient to propagate back across many time steps, neither exploding nor vanishing. In this sense, the gates are learning when to let an error in and when to let it out. The minimization of LSTM total error is achieved by implementing the iterative gradient descent such as backpropagation. It changes each weight in proportion to its derivative with respect to the error.

2.3 LSTM Implementation

The LSTM RNN model proposed in this study about, univariate case where only the SPI values are used. In this architecture, we use one layer of LSTM which consists of 100 cells,used relu as activation function.and next a dense layer is considered.trainable parameter for LSTM layer are 40,800 and for dense layer 101 and total parameter 40,901. For compile the model Adam optimizer is used were Adam optimizer gives much higher performance than the previously used and outperforms them by a big margin into giving an optimized gradient descent.Used loss function is MSE Here,I considered Training set as 603 data and Testing set as 201 data from total 804 dataset.

Chapter 3

Results and Discussions

In this work, I extracted 804 spi values with spi tool in R. After creating a dataset and applied lag with 5.Then loaded to LSTM.And considered Training set as 603 data and Testing set as 201 data from total 804 dataset.And predicted 201 data ,shown in figure 3.1 .When the predicted and actual output are plotted in x-y plane with respect to months and SPI values as shown in figure 3.2. which shows an efficient prediction of the model.The resultant model obtain an mean square error of 0.6492444328748223 as shown in figure 3.3,and obtained with root mean square error of 0.8057570557400179, mean absolute error of 0.6231234287528383 AS shown in figure 3.4

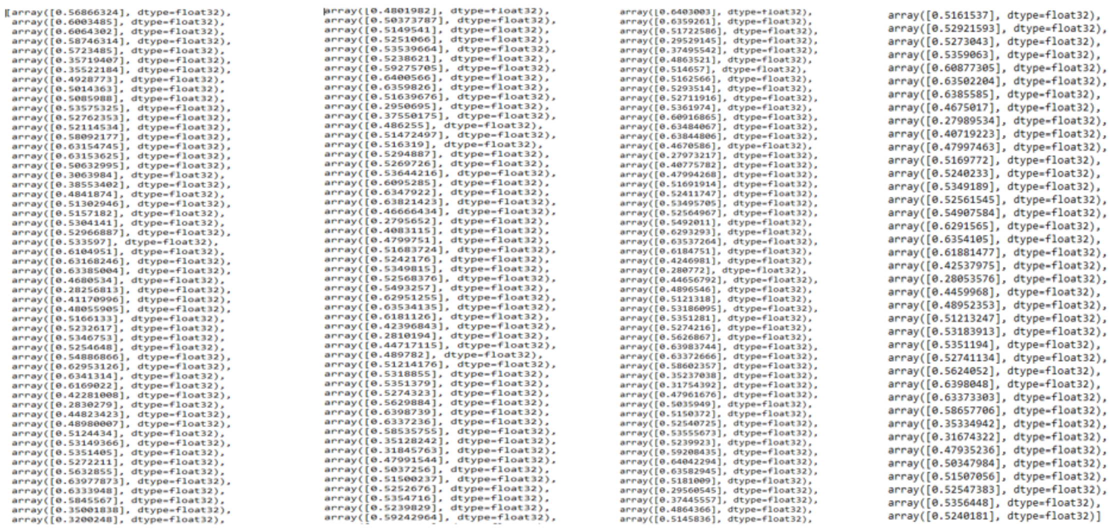

Figure 3.1: predicted 201 datas

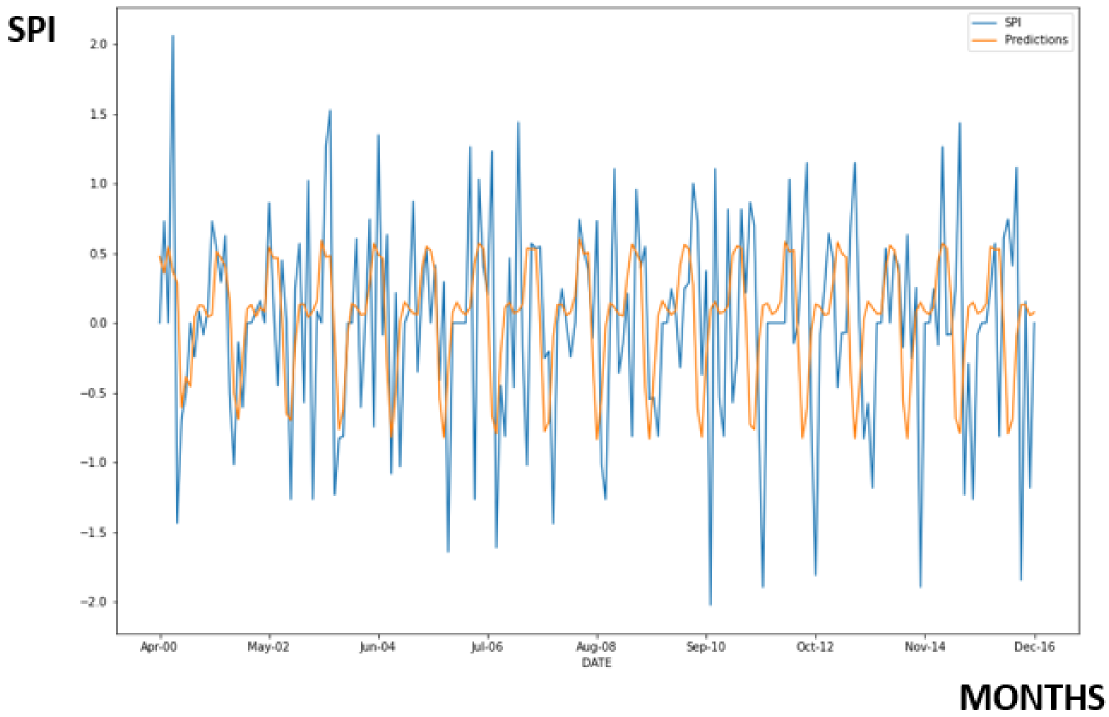

Figure 3.2: Comparison graph of actual and predicted outputs

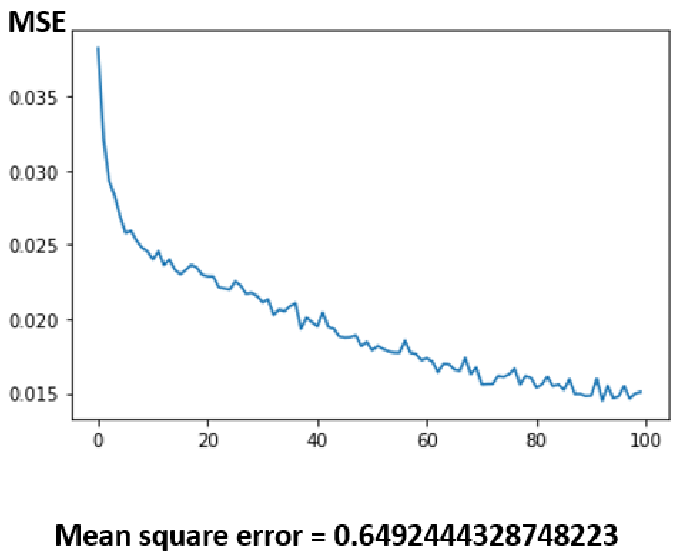

Figure 3.3: Loss function

Figure 3.4: MSE,RMSE,MAE

Chapter 4

Conclusion

Drought is one of the most Enviormental disaster that cause severe damage to nature and living things,Immediate drought impacts can include visibly dry vegetation and lower water levels in lakes and reservoirs. Longer-term impacts, such as land subsidence, seawater intrusion, and damage to ecosystems, can be harder to see, but more costly to manage in the future.and precipitation is random in character to it very complex to predict drought So, it is very essential to have a accurate predicting monitor system for taking security measure against drought.There are many statistical model available for the drought prediction but,it seem to be takes a lot time for the task. here we consider LSTM a deep learning model for predicting the drought and it shows an efficient output for the task with rainfall data alone.And also Exciting physical sensing or hydro-logical model are more data intensive and more parameters are used, we cannot compare or replace the AI model with the rainfall data alone.therefore,we can use the AI models as a complimentary tool for urgent tasks with rainfall data alone. Or by adding multivariate input like climatic oscillations,temperature,humidity the efficiency and accuracy of the model can be increased.Hence the model can be used for long-term

Bibliography

[1] Poornima S, Pushpalatha M. Drought prediction based on SPI and SPEI with varying timescales using LSTM recurrent neural network. Soft Computing. 2019 Sep;23(18):8399-412.

[2] X, Zhou J, Yu H, Liu D, Xie K, Chen Y, Hu J, Sun H, Xing F. The Development of a Hybrid Wavelet-ARIMA-LSTM Model for Precipitation Amounts and Drought Analysis. Atmosphere. 2021 Jan;12(1):74.

[3] W, Wu J, Cui H, Hu T. Drought Prediction Based on Feature-Based Transfer Learning and Time Series Imaging. IEEE Access. 2021 Jul 15;9:101454-68.

[4] Tian W, Wu J, Cui H, Hu T. Drought Prediction Based on Feature-Based Transfer Learning and Time Series Imaging. IEEE Access. 2021 Jul 15;9:101454-68.

[5] Zhang N, Dai X, Ehsan MA, Deksissa T. Development of a Drought Prediction System Based on Long Short-Term Memory Networks (LSTM). InInternational Symposium on Neural Networks 2020 Dec 4 (pp. 142-153). Springer, Cham.

[6] Sahoo BB, Jha R, Singh A, Kumar D. Long short-term memory (LSTM) recurrent neural network for low-flow hydrological time series forecasting. Acta Geophysica. 2019 Oct;67(5):1471-81.

[7] Dikshit A, Pradhan B. Interpretable and explainable AI (XAI) model for spatial drought prediction. Science of The Total Environment. 2021 Dec 20;801:149797.

[8] 8Dikshit A, Pradhan B. Explainable AI in drought forecasting. Machine Learning with Applications. 2021 Dec 15;6:100192.

[9] ZHANG, J.H., ZHANG, Q., XU, D.H. and DING, Y., Application of a combined ARIMA-LSTM model based on SPI for the forecast of drought: A case study in Qinghai Province. Arid Land Geography, 43(4), pp.1004-1013.