Drought Prediction with Feature Enhanced LSTM Model using Metaheuristic Optimization Algorithms

Leelavathy S. R., A. Mary Mekala*

SCOPE Vellore Institute of Technology Vellore, Tamil Nadu, India
Emails: leelavathys.r2015@vit.ac.in; amarymekala@vit.ac.in*;

Abstract

The impact of drought builds on all three fronts of economy, environment, and society is devastating. Predicting its arrival and duration is highly important to arrange any sort of mitigation plans. The association of detailed relationship between multiple variables makes drought prediction a highly complex task. Especially influence of global warming, polar sea extent variations and their influence on overall ocean temperature have altered the seasonal rainfall behaviors all over the world. In the midst of it, predictions centered on the history of rainfall levels become inaccurate. The proposed system is an optimized deep learning prediction model integrating indigenous knowledge (IK) is proposed to predict the drought. IK expressed in human language is translated using fuzzy function and fed to an improved Long Short Term Memory (LSTM) model. The LSTM model hyperparameters are optimized using a hybrid of Particle Swarm Optimization (PSO) with firefly to produce the meta-heuristics algorithm which will provide the best performance in presence of integration of IK features into modern meteorological features which solves the problem of local minima in LSTM hyperparameter optimization. The performance of the proposed results were tested compared with the meteorological information gathered by the Karnataka Natural Disaster Monitoring Centre (KNDMC) for the district named Chitradurga of the Karnataka state in India. The proposed system which is Indigenous Knowledge merged along the cross model attention network can produce at least 1.4% higher Nash–Sutcliffe model efficiency coefficient (NSE) and 30% lower Mean Absolute Error (MAE) in the prediction of Standard Precipitation Index (SPI) compared to Convolution Neural Networks (CNN) and LSTM based time series prediction models.

Keywords: Drought prediction; indigenous knowledge; fuzzy function; hyper parameter optimization; Long Short-Term Memory (LSTM); Firefly algorithm; Particle Swarm Optimization (PSO).

1. Introduction

The natural calamity drought having serious and large-scale impact on three fronts of environment, economy, and society. The impact is often long-term lasting for continuous months or even years [1] affecting food production, decreasing life probability and contracts economy. Over last three decades, increasing effects of global warming has affected earth's climate system. This has altered the level and the severity of droughts globally [2]. The raise in the past 20 years regarding the frequency of droughts is a key threat to human population and comprehensive knowledge of it is essential to prevent any catastrophes. The culmination of this comprehensive knowledge is drought prediction tool. The precision resulting drought prediction tool is vital to devise effective mitigation plans [3]. This is critical to nations like India, wherein agriculture comprises over 70 percent of the country's economic output is dependent on seasonal monsoon rainfall and drop in agricultural production poses huge food security and inflation problems to second largest populated county. Many approaches to prediction drought have already been put forth in earlier works. They forecast drought in the the Palmer Drought Severity Index (PDSI), as described by the authors in [4], among other types of indicators. The authors have defined the Standardized Precipitation Index (SPI) in [5], the Standardized Nonstationary Precipitation Index (SNSPI) in [6], the Joint Deficit Index (JDI) in [7], and the Copula-based Joint Drought Index [8]. Amongst them, SPI with a basic indication is among the most frequently employed techniques [9]. According to the precipitation index for previous times, it is computed. nevertheless recently many changes in precipitation amounts resulting from the rapid effects of global warming and extensive

Doi : https://doi.org/10.54216/JCIM.140208
Received: January 17, 2024 Revised: March 03, 2024 Accepted: July 03, 2024

environmental destruction, making precipitation predictions less accurate. Given this scenario, a great deal of effort has been conducted to improve the precision of drought predictions; one notable but still-developing methodology is the integrating of traditional knowledge into modern prediction systems.

In today's world, there is growing awareness about the value of traditional practices in drought prediction. A significant number of methods associated to traditional knowledge are strongly advised. But there are various advantages to using indigenous knowledge. Indigenous knowledge systems are significant. On the other hand, techniques to predict weather take into account a wider region. As a result, combining traditional knowledge systems with meteorological forecasting methods can increase the accuracy of predictions for specific regions and so lessen the risks to farmers.
There are multiple challenges in incorporating IK into modern prediction methods, though. Among them are (i) No standard procedures available for converting IK into the data forms required by recent prediction systems. (ii) There is no specified integration mechanism or technique for integrating the IK into mordent prediction systems. This paper introduces an enhanced deep learning model with fuzzy of IK for drought prediction, based on this previous findings. In order to translate the human language expressed IK into a deep learning feature and fuse it into an LSTM for drought prediction, this work used fuzzy translation.

Using a hybrid metaheuristics approach incorporating PSO and firefly, the LSTM hyper parameters are optimized to yield the best results when IK features are integrated with modern meteorological features. The contributions to the work are as follows. (i) A specialized fuzzy logic encoding method for IK feature encoding. (ii) A fusion metaheuristic approach was used to optimize the hyperparameters of the IK feature integrated LSTM model. This is how the rest of the paper is structured. The techniques used to optimize the built on both machine learning and deep learning algorithms for drought prediction will be deliberated in Subsection 2. Subsection 3 outlines in entirely the optimized a method for deep learning has been stated for drought prediction. The results as well as a description of the proposed deep learning mechanism discussed in Subsection 4. The conclusion and potential future scope of work are laid out in Subsection 5.

2. Literature Review

In recent days scenario the various techniques on deep learning have a huge impact on effective performance of the neural networks, the authors in [10] using artificial neural networks (ANNs) to forecast drought with one month lead time. To predict drought in terms of SPI, a parameter list which includes thirty atmospheric and soil-specific variables is provided into the ANN. Overfitting was avoided and reliability was improved using dropout and dense layer. In addition, a game theory-based technique called Shapely additive explanations is made use for estimation of average marginal contribution of feature to the prediction result and provide weightage to features.

Authors in [11] developed by making use of hydro meteorological variables as input, comprising temperature of air, the surface pressure, the speed of wind, soil moisture, vaporization, soil humidity, and geopotential height, a specific framework of deep learning model is trained to predict drought. A one-dimensional CNN was designed for forecasting drought in reference to standardized precipitation Anomaly Index (SPAI). Different combinations of following hyper parameters of: Filter size, number of filters, dropout percentage and number of hidden layers were experimented in different combinations and the best combination is selected. The configuration is sub optimal and search space is very limited. In [12] the authors recommended a new model as drought cast for predicting drought over a lead time of 1 -12 weeks. Meteorological variables giving inputs to a RNNs or recurrent neural networks to predict drought as a binary class of 0 or 1. Ensembling with dropout mechanism was employed to increase the forecast precision of the model.

The authors in [13] experimented with 16 diverse models of machine learning were employed. Last 90 days of 18 meteorological parameters as drought input parameters. Authors recommended effective Employing feature selection along with oversampling to boost precision in predictions.

The authors in [14] used deep forward neural network to predict agricultural drought from parameters of precipitation, vegetation, and soil factors. Drought was forecasted with respect to index soil moisture deficit index (SMDI). Author optimized the number of hidden layers and activation function through brute force to achieve the best configuration for deep forward neural network.

Doi : https://doi.org/10.54216/JCIM.140208
Received: January 17, 2024 Revised: March 03, 2024 Accepted: July 03, 2024

In [15] the authors combined functional connection between random vectors and various optimization algorithms to forecast drought using the scale of SPI. The Particle swarm optimization (PSO), the genetic algorithm (GA), grey wolf optimization (GWO), social spider optimization (SSO), salp swarm algorithm (SSA), and hunger games search algorithm (HGS) were experimented and HGS has been found to give improved efficiency than the other algorithms. The authors integrated the adaptive neuro-fuzzy inference system (ANFIS) with different meta-heuristic approaches such as particle swarm optimization (PSO), and ant colony algorithm (ACO), and finally genetic algorithm (GA), and butterfly optimization algorithm (BOA) to calculate drought as terms of SPI [16].

The authors in [17] used optimized algorithm artificial neural network (ANN) model for drought prediction using three metaheuristics algorithms namely the salp swarm algorithm (SSA), PSO, and GA. The best results were obtained in using SSA. In [18] the authors enhanced variational mode breakdown in the extreme learning machine and used it for drought prediction. The improvement was able to increase the efficiency of the drought estimate by at least 5%.

The authors in [19] used time series imaging to forecast drought in terms of SPEI. Feature based transfer learning is used to extract the features from SPEI image used as prediction features. Standard precipitation Evaporation Index (SPEI) sequences were converted to image using Recurrence plot. This work used transfer learning technique for enhancing the accuracy of predictions of a convolutional neural network model.

The authors in [20] for predicting drought determined by SPEI, the authors investigated with different combinations of various algorithms of machine learning. Extreme gradient Boost (EGB), random forest (RF), convolutional neural network (CNN), and long short-term memory (LSTM) are the four different machine learning models that the authors trained. There are two SPEI scales: SPEI-3, which predicts drought for three months, and SPEI-6, which predicts drought in six months. When comparing with various models, the LSTM model that had been trained with inputs consisting of rainfall, average temperature, lowest temperature, maximum temperature, and wind speed performed better. But hyper parameters of LSTM were kept at their default values without any optimization. The study of the literature demonstrates that almost all of the methods and models prediction of drought was based on the hydro meteorological parameters. Which caused lesser prediction accuracy and lead time. There were no provisions to involve cross domain information like IK into existing forecasting models. Though hyper parameter optimization is considered in many works, the optimization functions were not designed to minimize prediction error in presence of cross domain learning. Also, the optimization algorithms used in previous studies did not take into account over fitting and local minima problem while fine tuning the hyper parameters. This work considers these problems and proposes following solutions to the gaps in the existing works. i) Limited accuracy and lead time, the proposed Integrated IK with hydro meteorological parameters. Through this integration, accuracy was increased, and lead time was increased up to 6 months. ii) The optimization search space was limited. Did not consider over fit and local minima problem while optimizing the hyper parameters, the proposed solution used hybrid metaheuristics with many hyper parameters. Hybrid metaheuristics solves the local mining and over fitting problem. Many hyper parameters were considered for optimization to increase search space.

Table 1: Problems and solutions to the problems

Problems	**Solutions proposed in this work**
Limited accuracy and lead time	Integrated IK with hydro meteorological parameters. Through this integration, accuracy was increased and lead time was increased up to 6 months.
Search space for optimization was limited. Did not consider over fit and local minima problem while optimizing the hyper parameters	Used hybrid meta heuristics with many hyper parameters. Hybrid meta heuristics solves the local mining and over fitting problem. Many hyper parameters were considered for optimization to increase search space.

3. Methodology

Doi : https://doi.org/10.54216/JCIM.140208
Received: January 17, 2024 Revised: March 03, 2024 Accepted: July 03, 2024

Figure 1 illustrates the framework of the proposed method. The features that match the traditional knowledge specifications are encoded using fuzzy techniques. Coupled with meteorological variables that bring about the enhanced features, this is incorporated into the cross-model attention-based learning approach. The multivariate LSTM obtains the enriched features as an input, which facilitates the use of SPI in predicting the drought. The hybrid meta-heuristics algorithm is employed to adjust the fine-tuned LSTM hyperparameters. The following steps constitute the proposed solution: LSTM-based prediction, hyperparameter optimization, fuzzy, and cross-model attention learning below are insights on each step.

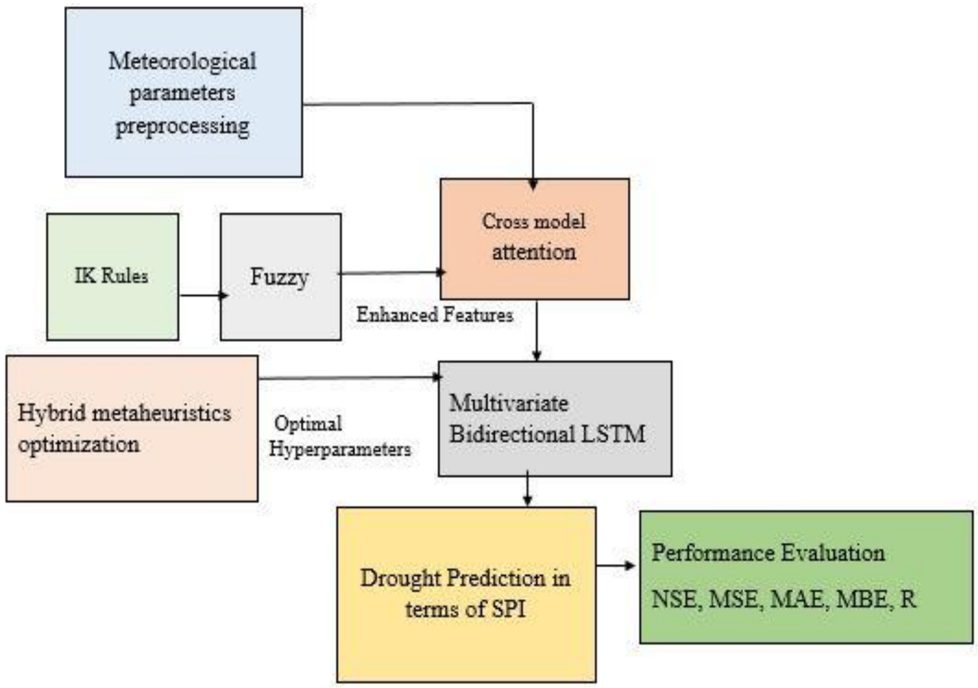


Figure 1: The proposed methodology of optimized deep learning technique

3.1. Fuzzy term weighting

i. Tokenization: Tokenize the text into individual terms as words or phrases. Ii. Term Frequency (TF) Calculation: Determine how often each term appears within the document. iii. Calculating the Inverse Document Frequency (IDF): determine the inverse document frequency for each term based on the entire document collection. Iv. Fuzzy Logic Operations: Define fuzzy logic operations (e.g., fuzzy AND, OR, NOT) to combine TF and IDF scores for each term. v. Weight Assignment: Apply fuzzy logic operations to assign weights to each term based on its TF and IDF scores. vi. Normalization: Normalize the weights to ensure they fall within a specific range (e.g., between 0 and 1).

3.2. Cross modal attention learning

The meteorological variables acquired from datasets and the meteorological variable values resulting from fuzzy are in two domains and must be fused using cross modal attention to obtain enhanced features is as shown in figure below Multi head attention proposed in is extended for multiple modalities attention in this work. Say there are two modalities$\{m_1, m_2\}$. The cross modal attention for modal m_1 takes output of its feature encoding layer as query vector and the output of m_2 feature encoding layers as key and value

Doi : https://doi.org/10.54216/JCIM.140208
Received: January 17, 2024 Revised: March 03, 2024 Accepted: July 03, 2024

vectors. It then applies multi head scaled dot product attention. It helps each modality to learn cross reference information from other modality. Finally features from the cross-modality attention of m_1 and m_2 are pooled using average fusion and an enhanced feature vector is generated.

.

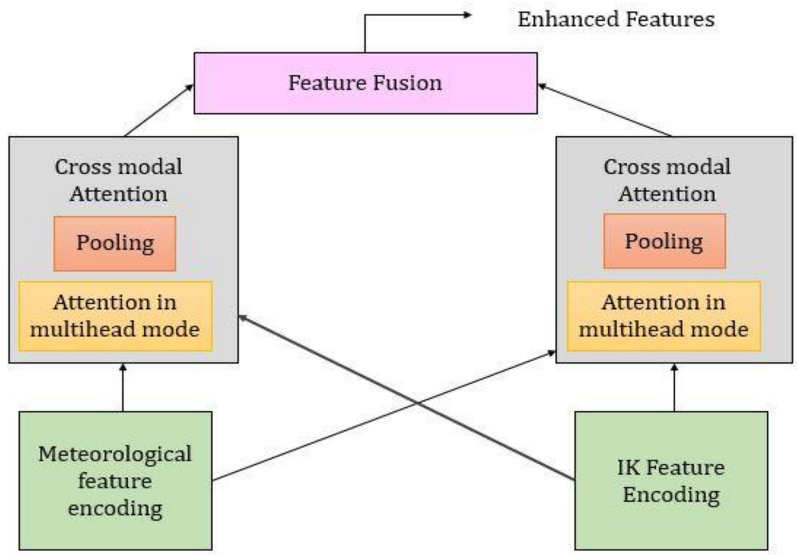


Figure 2: Cross modal attention learning

3.3. LSTM Based Prediction

The enhanced features are sequenced to a length N and this sequence is employed to train a multivariate LSTM to forecast drought with respect to of SPI value. LSTM is a more effective RNN (recurrent neural network) that has gating mechanism. This gating mechanism allows LSTM to retain or forget a level of information. This allows LSTM to control the learning rate. The input value for each LSTM cell comprises of the previously hidden state and the current input vector. A weighted sum of the inputs will be utilized for calculating the activation of the cell along with bias b. This receives an input and generates its hyperbolic tangent activation function.

$$c_t = \emptyset_t(W_c x_t + U_c h_{t-1} + b_c) \quad (1)$$

Where the terms c_t is the activation function of candidate cell. x_t is the input vector. W_C, U_c are corresponding weight matrices. h_{t-1} represents the hidden state vector with preceding time period and b_c representing bias. Gates control the data amount that will be remembered or forgotten. The level to be retained has been controlled by the input gate. The level of forgetting is set by the forget gate accessing hidden state data is the final gate.

$$f_t = \emptyset_s(W_f x_t + U_f h_{t-1} + b_f) \quad (2)$$

$$i_t = \emptyset_s(W_i x_t + U_i h_{t-1} + b_i) \quad (3)$$

$$o_t = \emptyset_s(W_o x_t + U_o h_{t-1} + b_o) \quad (4)$$

Where the terms f_t stands for the forgot gate vector. i_t $o_t$are the input and output gate vector respectively. LSTM takes the $z = (z_1, z_2, \dots z_t)$, the observed time t and time t+1 for prediction of the drought index and every z_i is the embedded input which is transformed as $x = (x_1, x_2, \dots x_t)$. In the context to regression, the Softmax classifier takes the final output of its

Doi : https://doi.org/10.54216/JCIM.140208
Received: January 17, 2024 Revised: March 03, 2024 Accepted: July 03, 2024

LSTM layer. Softmax classifier matching the LSTM output to one of the feasible values of the drought indicator in the regression initialization. Consider we have k drought indicators with the values {1, 2…k}. The possibility for each of the k values must be predicted by the softmax classifier. The k-dimensional vectors comprising the k predicted possibilities is the outcome of the softmax classifier. The loss function which is utilized to train the softmax regression classifier:

$$L = -[\sum_{i=1}^{m} \sum_{k=0}^{1} 1\{y^{(i)} = k\} \log P(y^{(i)} = k|z^{(i)}; \theta)] \quad (5)$$

Where L is the loss function, P is the exponential function, k is estimated probabilities

$$P\left(y^{(i)} = k \middle| z^{(i)}; \theta\right) = \frac{\exp(\theta^{(k)} z^{(i)})}{\sum_{j=1}^{K} \exp(\theta^{(k)} z^{(i)})}$$

Where $\theta^{(1)}$,$\theta^{(2)}$,…$\theta^{(k)}$ are model parameters and $\exp(\theta^{(k)} z^{(i)})$ is the feature vector value for normalization of parameter.

3.4. Hyper parameter optimization

By optimizing the LSTM's hyperparameters, one may enhance the model's ability to predict drought. Table 2 specifies various ranges and hyperparameters.

Table 2: Hyper-parameter to be fine-tuned

Variable	Specified Hyperparameter	Specified Range
P1	num_kernels	[16, 24, 32, 40, 48, 52,64]
P2	kernel_columns	[2-11]
P3	kernel_rows	[2-11]
P4	stride_x	[1-4]
P5	stride_y	[1-4]
P6	fullyconnected_units	[50-150]
P7	batch_size	[16, 24, 32, 40, 48, 52,64]

It is difficult to try the LSTM configuration in all possible combinations of the seven hyper parameters given in Table 2. The search space is huge, and a heuristics solution can be obtained using optimization algorithms. Usage of single optimization algorithms has higher probability of solution hitting the local minima. This work attempts to address this issue using a hybrid metaheuristic which combines the algorithms to optimization using additional characteristics of exploration and exploitation. The Pseudo code for hybrid metaheuristics algorithm is as shown below:

Pseudo code for Hybrid metaheuristics algorithm

```
//Initial solution using firefly
Population=[]
for i=1: population_size
    for j=1: solution size
```


Doi : https://doi.org/10.54216/JCIM.140208
Received: January 17, 2024 Revised: March 03, 2024 Accepted: July 03, 2024

```
            sol[j] = random values for variables P1 to P7 as specified  in the Table 2
            Population [i]. Sol.append (sol[j])
    end
end
For i=1:population_size
  Population[i]. Fitness= Maximum of fitness for solutions in the Population[i]
end
 While (iteration<=max_iteration)
        For  i=1: population_size
                   For j=1:population_size
                            If Population[i].fitness > Population[j].fitness
                                     Copy random Population[i] solutions towards Population[i].solutions
                  end
                  Population[i]. Fitness= Maximum of fitness for solutions in the Population[i]
        end
        Iteration++
  end
  BestPop← Population with maximum fitness
 // Taking the Best Sol from firefly, PSO algorithm is started
 For j=1: solution_size
      Particle[j] ← BestPop.Sol[j]
end
Iteration=1
Do
     Gbest←0
    For each Particle i
           F←Calculate Fitness using Equation (11)
           If F >  Pbesti
               Pbesti← F
           end
            If Pbesti>Gbest
                    Gbest←Pbesti
           end
     end
     For each Particle i
                 Update velocity based on Gbest
                  For each Velocity value
                      Update Particle[i].Sol to Gbest[i] Sol
                  end
     end
     Iteration++
While maximum iteration or minimum error criteria are not met
Return Gbest
```

Figure 3 shows metaheuristic algorithm for optimization.PSO has been employed in the study with exploration capability is combined with firefly algorithm with exploitation capability to obtain the optimization values for seven hyper parameters. Social behavior of swarms is the fundamental logic of the PSO swarm intelligence algorithm as stated by in reference [26]. Features like simplicity, flexibility and versatility make PSO most adopted compared to existing algorithms. In PSO the candidate solution to a problem is mapped to particle. The particles travel at various speeds and varied velocities and update their positions. The particles fine tune its velocity based on local and global best (p_{best}, g_{best}) values. The position of the particles is influenced by its current position, velocity, distance to local best and global best. The motion of particles occurs in iteration and in each iteration

Doi : https://doi.org/10.54216/JCIM.140208
Received: January 17, 2024 Revised: March 03, 2024 Accepted: July 03, 2024

The particles advance in the direction that is of the finest possible solution by exploring all solutions within the search space. For a particle X, it current position X_i and current velocity V_i is simplified as

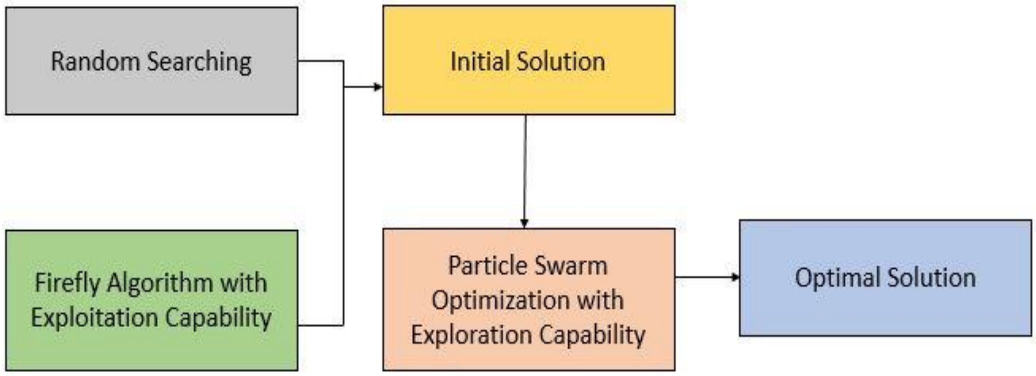


Figure 3: Hybrid Metaheuristic Algorithm for Optimization

$$X_i(t+1) = X_i(t) + V_i(t+1) \quad (6)$$

$$V_i(t+1) = wV_i(t) + c_1 r_1 \left(p_{besti}(t) - X_i(t)\right) + c_2 r_2 \left(g_{besti}(t) - X_i(t)\right) \quad (7)$$

In the above equations 6 and 7, t is the iterative value. c_1 and c_2 are acceleration coefficients, r_1 and r_2 are random numbers, w is the inertia weight. The iteration stops when no modification in the position of the particle or maximum iteration count has reached. Like PSO the authors in [27] use firefly as swarm intelligence algorithm modeling firefly natural behavior. Fireflies emit flashes, it emits flashes for various purposes like mating alternating etc. Based on the flashes the firefly moment is decided. The intensities of light attractiveness level affects the motion of fireflies. Light intensity I is calculated as inverse of square of distance r between the emitter and observer firefly.

$$I = \frac{1}{r^2} \quad (8)$$

Attractiveness level is in direct proportion to intensity of light is as shown

$$\beta(r) = \beta_0 \, e^{-\gamma r^2} \quad (9)$$

At r=0, the attractiveness is expressed as β_0, where γ is the light absorption coefficient. R is computed using the Euclidean distance formula. Between any two fireflies (i, j) , the movement of i (FX_i) is influenced by the attraction from j (FX_j). It is evaluated as shown below in equation 10.

$$FX_i = FX_i + \beta_0 e^{-\gamma r^2}(FX_j - FX_i) + \alpha \epsilon_i \quad (10)$$

The randomization parameter α is as stated in equation number 10 and ϵ_i is a random number in the hyper parameter optimization process, firefly algorithm starts first with initial random values for seven hyperparmeters. The firefly algorithm does better exploitation and results in initial optimal solution. PSO algorithm takes the initial optimal solution returned by firefly is the input given and offers the final optimal as output. The LSTM is configured with the optimal solution returned by PSO and utilized for training and subsequent classification the fitness function for optimization for both firefly and LSTM is designed to minimize the prediction error as

Doi : https://doi.org/10.54216/JCIM.140208
Received: January 17, 2024 Revised: March 03, 2024 Accepted: July 03, 2024

$$F = \frac{1}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(P_i - A_i)^2}} \quad (11)$$

In this scenario, A_i represents the actual SPI value, P_i is the SPI value that the LSTM predicted, and n is the test set size.

3.5 Data sets

Chitradurga district spans across 6 taluks: Chitradurga, Hiriyur, Hosadurga, Holalkere, Challakere, and Molakalmuru. It is situated in the hills of the Deccan Plateau in the valley of the Vedavati River, 202 km from Bengaluru. Around 66.75% of the district's people are employed in agriculture, either directly as farmers or indirectly as labour in the field. Agriculture is the district's primary source of income.

The district's significant agricultural products are coconut, pomegranate, groundnut, onion, ragi, and maize. Covering 15°00' to 16° 30' North latitude and 75°40' to 77° 0' East longitude, the district of the region has a over-all area of 10.78 lakh hectares. This area experiences 58–76% mean humidity and 592.5 mm of precipitation yearly on average. The area experiences 32 days of precipitation year, particularly June through September being the typical wet season. The majority of this region's rainfall is caused by the southwest monsoon season. Meteorological data from 1987 to 2017 was provided by the Karnataka State Natural Disaster Monitoring Centre (KSNDMC) to the Chitradurga district.

The temperature, humidity, wind speed, and precipitation data for an area in particular over a number of years at regular intervals are incorporated in the dataset. On a monthly basis, the SPI value is measured. All series of temperature, humidity, also the wind speed has a chosen output SPI value. Collectively these series can be supplied as input sequences to be combined at three- and six-monthly time frames, correspondingly. The remaining eighty percent of the dataset used for training, while twenty percent is employed for testing. The retrieved IK rules from [23] have been applied for fuzzy functions.

4. Results

Using the Keras and Tensorflow modules, the Indigenous Knowledge integration methodology was implemented in Python 3.6. The Karnataka district of Chitradurga has been selected to test the IK integration since it encompasses documented meteorological observations in addition to traditional knowledge systems.

The dataset was structured in form of temperature (minimum and maximum), humidity (minimum and maximum), wind speed (minimum and maximum), precipitation (minimum and maximum) in interval of months. The Standardized Precipitation Index values for interval of 3 months and 6 months were collected and set as output and the sequence of meteorological observations in those 3 months and 6 months were set as input and a processed dataset of SPI-3 and SPI-6 were generated. Two different IK fused LSTM models were trained with SPI-3 and SPI-6 using the datasets in 80:20 ratio for training and testing. The IK rules extracted from [26] were used for fuzzy transformation.

The effectiveness of drought prediction is measured in terms of following parameters of Nash–Sutcliffe model Efficiency Coefficient (NSE) [22], the performance parameter mean square error (MSE), the performance parameter mean absolute error (MAE), the performance parameter mean bias error (MBE) and the parameter prediction correlation coefficient (R). All the parameters are calculated correlating actual SPI value (A_i) and predicted SPI value (P_i) as below

$$NSE = 1 - \frac{\sum(P_i - A_i)^2}{\sum(\bar{A} - A_i)^2} \quad (12)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|A_i - P_i| \quad (13)$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(P_i - A_i)^2 \quad (14)$$

$$MBE = \frac{1}{n}\sum_{i=1}^{n}(P_i - A_i) \quad (15)$$

Doi : https://doi.org/10.54216/JCIM.140208
Received: January 17, 2024 Revised: March 03, 2024 Accepted: July 03, 2024

$$R = \frac{\sum_{i=1}^{n}(\bar{A}-A_i)((\bar{P}-P_i)}{\sqrt{\sum_{i=1}^{n}(\bar{A}-A_i)^2 \sum_{i=1}^{n}(\bar{P}-P_i)^2}} \quad (16)$$

In the above equations of the metrics NSE, MAE, MSE, MBE and R, the value n is the total number of observation after testing, A is the actual value and P is the value that is predicted. The performance of proposed optimized deep learning model is compared against CNN based prediction with enhanced features found by cross model learning (C1), LSTM trained with meteorological variables alone (C2) and LSTM without optimization (C3). The CNN and LSTM models proposed in works of [20] were used for realization of C1, C2 and C3. Six scenarios with different combination of input meteorological variables (rainfall, relative humidity, wind speed, average, minimum, and maximum temperature, as well as sunshine) were created.

The performance of the results (Proposed, C1, C2, and C3) were tested in these seven scenarios for SPI-3 and SPI-6 prediction outputs. The input combination for six scenarios is listed in Table 3.

Table 3: Test scenarios

Scenario	Input parameters
S1	Rain, average temperature
S2	Rain , average temperature, max temperature
S3	Rain, average temperature, max temperature, min temperature
S4	Rain , average temperature, max and min temperature, wind speed
S5	Rain , average temperature, max and min temperature, wind speed, relative humidity
S6	Rain, average,max,min temperature wind speed, relative humidity, sunshine

The results for Mean Square Error (MSE) and Nash–Sutcliffe model Efficiency Coefficient (NSE) are given below table 4

Table 4: Comparison of MSE and NSE

	MSE				**NSE**			
Scenario	Proposed optimized model	C1 CNN based model	C2 LSTM with meteorological variables	C3 LSTM without optimization	Proposed optimized model	C1 CNN based model	C2 LSTM with meteorological variables	C3 LSTM without optimization
				SPI-3				
1	0.11	0.17	0.18	0.14	0.72	0.68	0.68	0.72
2	0.09	0.15	0.14	0.13	0.72	0.68	0.68	0.72
3	0.08	0.11	0.12	0.11	0.73	0.69	0.69	0.73
4	0.07	0.10	0.11	0.10	0.75	0.69	0.69	0.73
5	0.08	0.10	0.13	0.12	0.77	0.70	0.69	0.74
6	0.07	0.10	0.12	0.12	0.78	0.71	0.71	0.75
				SPI-6				
1	0.09	0.16	0.17	0.15	0.76	0.68	0.68	0.69
2	0.09	0.14	0.14	0.12	0.76	0.68	0.68	0.70
3	0.09	0.10	0.12	0.11	0.76	0.69	0.69	0.71
4	0.09	0.10	0.12	0.10	0.77	0.70	0.70	0.71
5	0.08	0.10	0.12	0.12	0.79	0.71	0.71	0.74
6	0.04	0.10	0.11	0.12	0.80	0.74	0.74	0.75

Doi : https://doi.org/10.54216/JCIM.140208
Received: January 17, 2024 Revised: March 03, 2024 Accepted: July 03, 2024

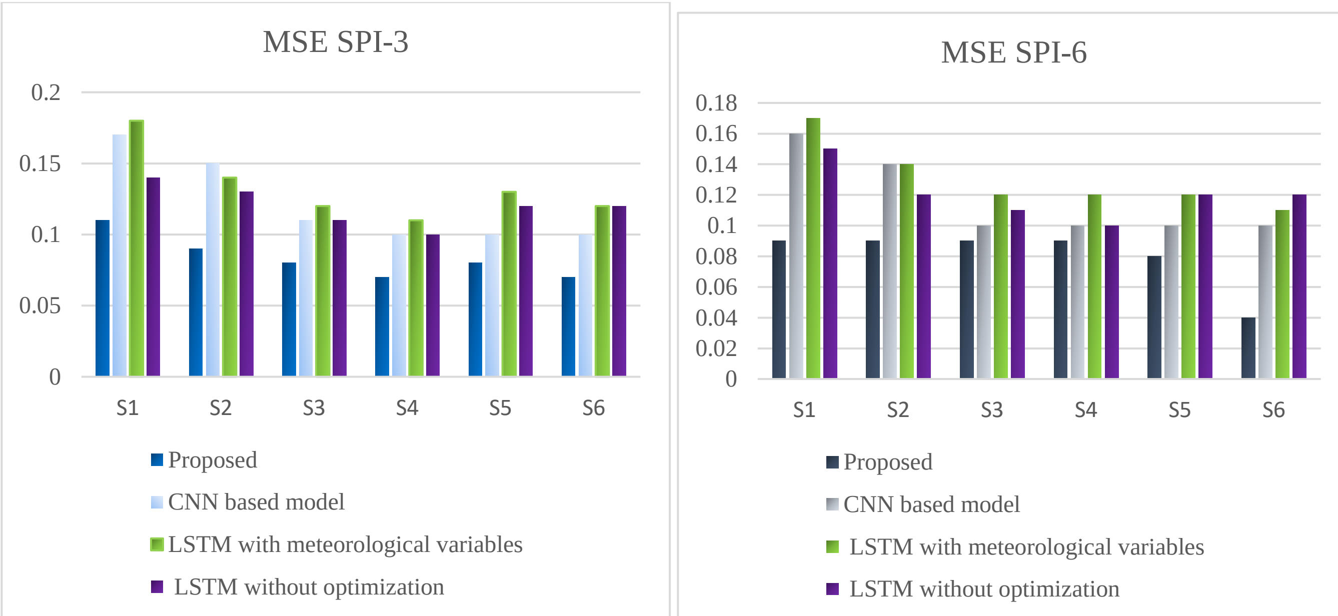


Figure 4: Comparison of MSE

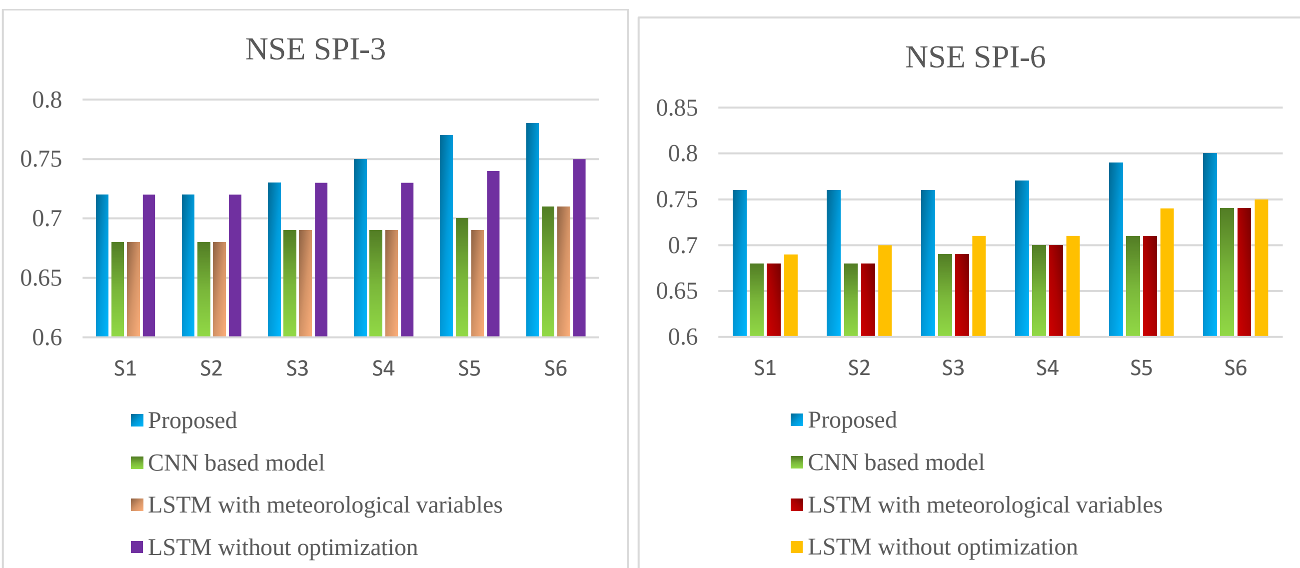


Figure 5: Comparison of NSE

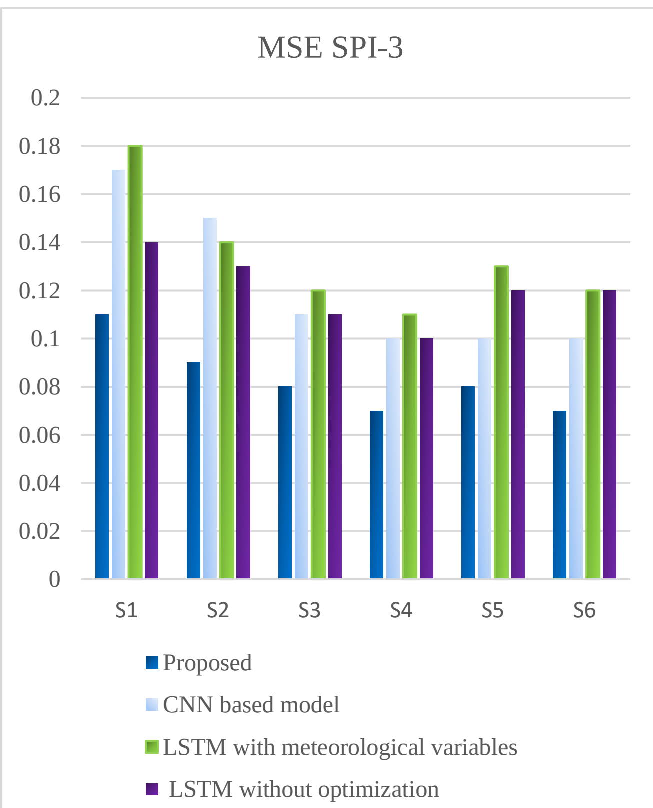
The results for Mean Absolute Error (MAE) and Mean Bias Error (MBE) are given below table 5

Table 5: Comparison of MAE and MBE

	MAE	**MBE**

Doi : https://doi.org/10.54216/JCIM.140208
Received: January 17, 2024 Revised: March 03, 2024 Accepted: July 03, 2024

Scenario	Proposed optimized model	C1 CNN based model	C2 LSTM with meteorological variables	C3 LSTM without optimization	Proposed optimized model	C1 CNN based model	C2 LSTM with meteorological variables	C3 LSTM without optimization
SPI-3								
1	0.36	0.41	0.41	0.36	-0.20	0.11	0.12	-0.24
2	0.36	0.40	0.40	0.36	-0.20	0.11	0.11	-0.24
3	0.35	0.39	0.40	0.35	-0.20	0.11	0.11	-0.23
4	0.34	0.38	0.39	0.34	-0.20	0.10	0.10	-0.22
5	0.32	0.37	0.38	0.33	-0.20	0.08	0.08	-0.21
6	0.31	0.36	0.37	0.32	-0.18	0.07	0.07	-0.18
SPI-6								
1	0.32	0.39	0.38	0.32	-0.026	-0.05	0.07	-0.13
2	0.31	0.38	0.38	0.32	-0.026	-0.05	0.07	-0.13
3	0.30	0.38	0.38	0.31	-0.026	-0.05	0.06	-0.12
4	0.29	0.37	0.38	0.30	-0.026	-0.04	0.06	-0.11
5	0.28	0.36	0.37	0.29	-0.025	-0.03	0.05	-0.10
6	0.27	0.35	0.36	0.28	-0.024	-0.02	0.04	-0.09



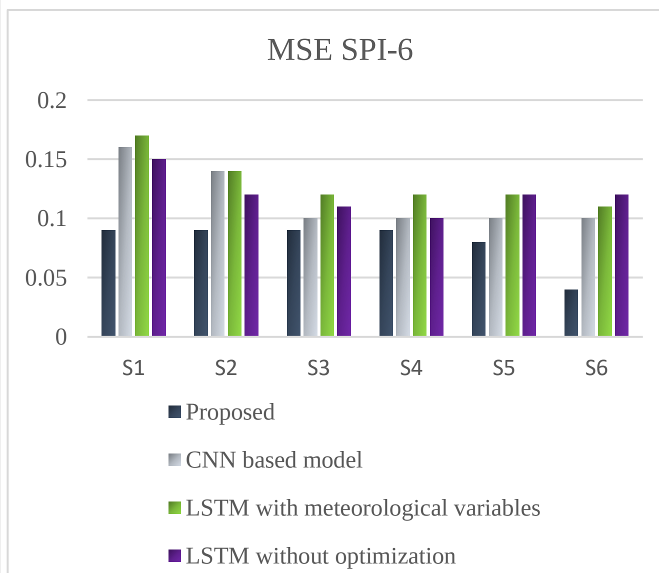


Figure 6: Comparison of MAE

Doi : https://doi.org/10.54216/JCIM.140208
Received: January 17, 2024 Revised: March 03, 2024 Accepted: July 03, 2024

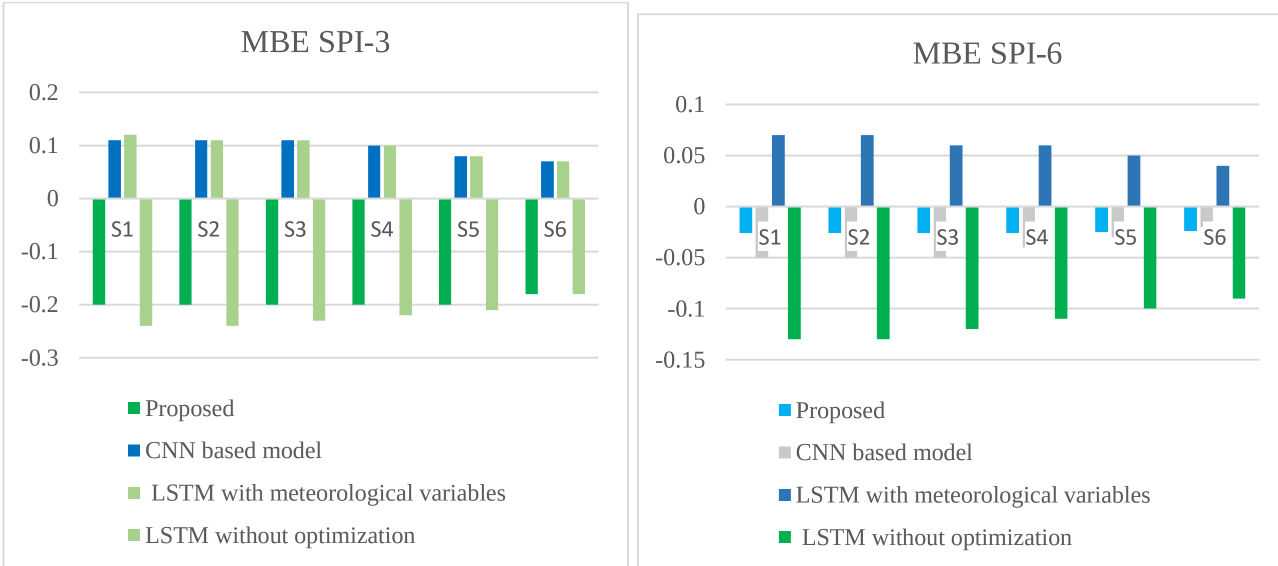


Figure 7: Comparison of MBE

The results for R value are given below table 6.

Table 6: Comparison of R

	R			
Scenario	Proposed optimized model	C1 CNN based model	C2 LSTM with meteorological variables	C3 LSTM without optimization
		SPI-3		
1	0.82	0.77	0.76	0.80
2	0.83	0.78	0.77	0.81
3	0.84	0.79	0.78	0.82
4	0.85	0.80	0.79	0.83
5	0.86	0.81	0.80	0.84
6	0.87	0.82	0.81	0.85
		SPI-6		
1	0.82	0.78	0.78	0.81
2	0.83	0.79	0.79	0.82
3	0.84	0.80	0.80	0.83
4	0.85	0.81	0.81	0.84
5	0.86	0.82	0.82	0.85
6	0.87	0.83	0.83	0.86

Doi : https://doi.org/10.54216/JCIM.140208
Received: January 17, 2024 Revised: March 03, 2024 Accepted: July 03, 2024

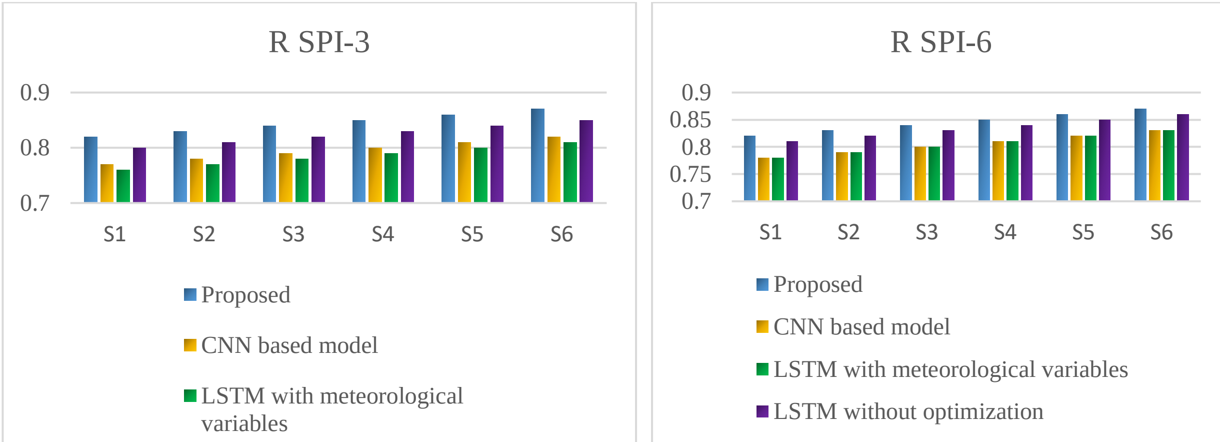


Figure 8: Comparison of R

Figures 4,5, 6, 7, and 8 illustrate the comparisons of MSE, NSE, MAE, MBE, and R. Table 4 depicts the comparison of the mean square error (MSE) for every solution over all SPI scales. The prediction accuracy is higher when the MSE value is lower. In comparison to existing work, the proposed solution has an MSE that is 64% lower. Table 5 summarizes the results of comparing the NSE across all SPI scale solutions. The better the prediction, the higher the NSE value. The NSE of the suggested solution is 8% higher than that of the current work. Table 6 illustrates the results of comparing the MAE across all solutions for all SPI scales. Improved prediction accuracy can be determined by a lower MAE number. The proposed strategy has a 34% lower MAE than previous works. Table 7 illustrates the results of examining the MBE across all solutions for all SPI scales. In comparison to LSTM, the proposed strategy has a 36% lower MBE. Table 8 presents the outcome of examining the R across all solutions for the SPI scales. Better prediction accuracy is indicated by a higher R value. In comparison to previous studies, the proposed solution has a 5.4% higher R value. As an outcome, the approach proposed outperforms the CNN and LSTM models that the authors suggested in [20]. The contribution of every component listed in Table 9 to the suggested solution's prediction performance (measured in terms of the MSE of the SPI value) has been investigated using ablation. Figure 9 presents the ablation results.

Table 7: Ablation tests

Ablation 1	Metrological variables + LSTM
Ablation 2	IK fused Metrological variables + LSTM
Ablation 3	IK fused Metrological variables + LSTM + Hyper parameter optimization

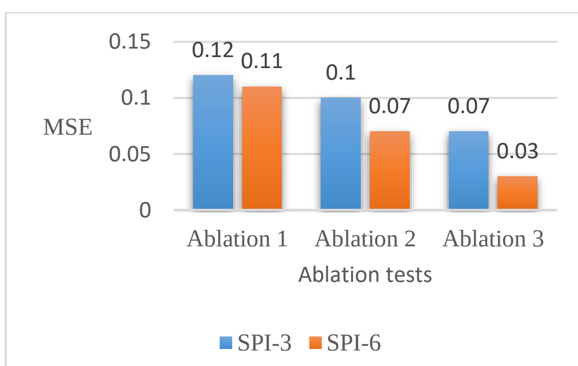


Figure 9: Ablation Results

Doi : https://doi.org/10.54216/JCIM.140208
Received: January 17, 2024 Revised: March 03, 2024 Accepted: July 03, 2024

The MSE in ablation 1 is 2.3 times higher compared to ablation 2 and MSE in ablation 2 is 1.7 times higher compared to ablation 3. Thus optimization reduces the MSE by around 0.6 times and IK fusion reduces the MSE by 1.35 times. This is in-line with the results shown in Table 4, where data based changes (cross model learning) is able to achieve lower MSE compared to LSTM model optimization. Thus fusing more IK rules can be considered to further improve the drought prediction accuracy in future work. The Friedman test was implemented to rank the various optimization approaches, and the outcome is presented in Table 10.

Table 8: Friedman test ranking results

Method used	Tested Friedman test rank	Obtained Final Rank
Proposed Particle Swarm Optimization +Firefly	1.320	Rank 1
Firefly only	2.900	Rank 2
Particle Swarm Optimization only	4.450	Rank 4

The Friedman test ranking is given by

$$Fr = \frac{12}{nk(k+1)}\left(\sum_{j=1}^{k}(R^2{}_j)\right) - 3n(k+1) \tag{18}$$

Where F_r is the Friedman test rank, n is the sample size, j is the measurement time, k the number of measurement times R_j the sum of ranks at measurement time j. Table 12 shows the Friedman rank as 1.320, 2.900, and4.450 for the proposed PSO+Firefly, firefly, PSO. Lower the value better is the ranking.

5. Conclusion

By integrating traditional knowledge systems with meteorological forecasting techniques, native forecasting precision can be enhanced and farmer risks may be reduced. Still, there have been many challenges that get to overcome while integrating IK into modern prediction techniques. For example, there are no standard approaches for transforming IK into the data forms necessary by contemporary prediction techniques, and there is no specified integration mechanism or procedure for integrating IK into more sophisticated prediction tools. This investigation provided a more effective deep-learning approach for drought prediction which incorporates indigenous knowledge's fuzzy function. Texts containing traditional knowledge have been transformed into meteorological parameters. Using cross-model attention learning, these characteristics are combined with historical meteorological variables to enhance the combination. Subsequently, the LSTM model is trained with the enhanced characteristics, and its hyperparameters have been optimized through a hybrid meta-heuristics algorithm which combines Particle Swarm Optimization (PSO) and firefly. This method yields optimal performance when IK features are integrated with contemporary meteorological features. The proposed solution's performance has been assessed using meteorological data from the Karnataka district of Chitradurga. The proposed approach outperformed CNN and LSTM-based models suggested in previous investigations in terms of Standard Precipitation Index (SPI) prediction precision, with an NSE value greater by at least 1.4% and an MAE lower by at least 30%.

References

[1] Pfister C, Weingartner R, Luterbacher J. Hydrological winter droughts over the last 450 years in the Upper Rhine basin: a methodological approach. Hydrological sciences journal. 2006 Oct 1;51(5):966-85.

[2] Hemamalini, Selvamani, and Visvam Devadoss Ambeth Kumar. (2022). Outlier Based Skimpy Regularization Fuzzy Clustering Algorithm for Diabetic Retinopathy Image Segmentation. Symmetry, 14(12), 2512

Doi : https://doi.org/10.54216/JCIM.140208
Received: January 17, 2024 Revised: March 03, 2024 Accepted: July 03, 2024

[3] Sathya Preiya, V., and V. D. Ambeth Kumar. (2023). Deep Learning-Based Classification and Feature Extraction for Predicting Pathogenesis of Foot Ulcers in Patients with Diabetes. Diagnostics 13(12), 1983.

[4] Guttman NB. A sensitivity analysis of the palmer hydrologic drought index 1. JAWRA Journal of the American Water Resources Association. 1991 Oct;27(5):797-807.

[5] McKee TB, Doesken NJ, Kleist J. The relationship of drought frequency and duration to time scales. InProceedings of the 8th Conference on Applied Climatology 1993 Jan 17 (Vol. 17, No. 22, pp. 179-183).

[6] Russo S, Dosio A, Sterl A, Barbosa P, Vogt J. Projection of occurrence of extreme dry-wet years and seasons in Europe with stationary and nonstationary Standardized Precipitation Indices. Journal of Geophysical Research: Atmospheres. 2013 Jul 27;118(14):7628-39.

[7] Kao SC, Govindaraju RS. A copula-based joint deficit index for droughts. Journal of Hydrology. 2010 Jan 15;380(1-2):121-34.

[8] Won J, Choi J, Lee O, Kim S. Copula-based Joint Drought Index using SPI and EDDI and its application to climate change. Science of the Total Environment. 2020 Nov 20;744:140701.

[9] Amrit K, Pandey RP, Mishra SK. Characteristics of meteorological droughts in northwestern India. Natural Hazards. 2018 Nov;94:561-82.

[10] Felsche E, Ludwig R. Applying machine learning for drought prediction in a perfect model framework using data from a large ensemble of climate simulations. Natural Hazards and Earth System Sciences. 2021 Dec 3;21(12):3679-91.

[11] Balakrishnan, Chitra, and V. D. Ambeth Kumar. (2023). IoT-Enabled Classification of Echocardiogram Images for Cardiovascular Disease Risk Prediction with Pre-Trained Recurrent Convolutional Neural Networks. Diagnostics 13(4), 775

[12] Brust C, Kimball JS, Maneta MP, Jencso K, Reichle RH. Droughtcast: a machine learning forecast of the United States drought monitor. Frontiers in big Data. 2021 Dec 21;4:773478.

[13] Jiang W, Luo J. An evaluation of machine learning and deep learning models for drought prediction using weather data. Journal of Intelligent & Fuzzy Systems. 2022 Jan 1;43(3):3611-26.

[14] Prodhan FA, Zhang J, Yao F, Shi L, Pangali Sharma TP, Zhang D, Cao D, Zheng M, Ahmed N, Mohana HP. Deep learning for monitoring agricultural drought in South Asia using remote sensing data. Remote Sensing. 2021 Apr 28;13(9):1715.

[15] Adnan RM, Mostafa RR, Islam AR, Gorgij AD, Kuriqi A, Kisi O. Improving drought modeling using hybrid random vector functional link methods. Water. 2021 Dec 1;13(23):3379.

[16] Kisi O, Gorgij AD, Zounemat-Kermani M, Mahdavi-Meymand A, Kim S. Drought forecasting using novel heuristic methods in a semi-arid environment. Journal of Hydrology. 2019 Nov 1;578:124053.

[17] Banadkooki FB, Singh VP, Ehteram M. Multi-timescale drought prediction using new hybrid artificial neural network models. Natural Hazards. 2021 Apr;106:2461-78.

[18] Liu Y, Wang LH, Yang LB, Liu XM. Drought prediction based on an improved VMD-OS-QR-ELM model. Plos one. 2022 Jan 6;17(1):e0262329.

[19] 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.

[20] Mokhtar A, Jalali M, He H, Al-Ansari N, Elbeltagi A, Alsafadi K, Abdo HG, Sammen SS, Gyasi-Agyei Y, Rodrigo-Comino J. Estimation of SPEI meteorological drought using machine learning algorithms. IEEE Access. 2021 Apr 20;9:65503-23.

Doi : https://doi.org/10.54216/JCIM.140208
Received: January 17, 2024 Revised: March 03, 2024 Accepted: July 03, 2024

[21] Ambeth Kumar, V.D. Ramakrishnan,M. (2013). Temple and Maternity Ward Security using FPRS. Journal of Electrical Engineering & Technology, 8(3), 633-637.

[22] Nash JE, Sutcliffe JV. River flow forecasting through conceptual models part I—A discussion of principles. Journal of hydrology. 1970 Apr 1;10(3):282-90. Liang JJ, Qu BY, Suganthan PN. Problem definitions and evaluation criteria for the CEC 2014 special session and competition on single objective real-parameter numerical optimization. Computational Intelligence Laboratory, Zhengzhou University, Zhengzhou China and Technical Report, Nanyang Technological University, Singapore. 2013 Dec;635(2):2014.

[23] Abdel Nasser H. Zaied, Mahmoud Ismail and Salwa El- Sayed, A Survey on Meta-heuristic Algorithms for Global Optimization Problems, Journal of Intelligent Systems and Internet of Things,Volume 1 , Issue 1 , PP: 48-60, 2020

Doi : https://doi.org/10.54216/JCIM.140208
Received: January 17, 2024 Revised: March 03, 2024 Accepted: July 03, 2024