JSR Publications

Available Online

J. Sci. Res. **17** (1), 71-85 (2025)

JOURNAL OF SCIENTIFIC RESEARCH

www.banglajol.info/index.php/JSR

IoT Aware Smart Agriculture Using Extreme Learning Classifier Based Predictive Analytics in a Cloud Environment

R. Sangeetha[1*], N. Girubagari[1], M. Ellakkiya[2]

[1]PG & Research Department of Computer Science, Cauvery College for Women (Autonomous), Affiliated to Bharathidasan University, Tiruchirappalli, 620 018, India

[2]Department of Computer Applications, Cauvery College for Women (Autonomous), Affiliated to Bharathidasan University, Tiruchirappalli, 620 018, India

Received 5 May 2024, accepted in final revised form 19 October 2024

Abstract

Agriculture encompasses soil nurturing, crop cultivation, and influencing human life and the environment for global economic growth. Effective irrigation and soil moisture management directly impact crop yields. To optimize productivity, an IoT-based soil monitoring system analyses soil parameters and weather conditions, generating substantial data stored on cloud platforms for predictive analytics. However, traditional methods face challenges in accurate and timely predictions. Addressing this, a novel Proximity Scaling Laplace Kernelized Extreme Learning Classifier (PSLKELC) is proposed. It consists of three stages: preprocessing, dimensionality reduction, and classification. Preprocessing involves data cleaning and transformation. Dimensionality reduction utilizes McNemar statistic multidimensional scaling to select relevant variables. Finally, the Laplace kernelized Extreme Learning classifier predicts soil moisture using the reduced dataset. The experimental evaluation compares the PSLKELC method with conventional techniques, considering metrics like accuracy, mean absolute error, time, and space complexity across various data sample sizes. Results demonstrate that PSLKELC enhances soil moisture prediction accuracy with reduced time and space complexity compared to traditional methods.

Keywords: Soil moisture prediction; Agriculture; Cloud; IoT.

doi: https://dx.doi.org/10.3329/jsr.v17i1.72958 J. Sci. Res. **17** (1), 71-85 (2025)

1. Introduction

Integration of Internet of Things (IoT) technology and cloud-based systems is an efficient approach to improving crop productivity by enabling farmers to monitor and manage soil moisture levels, make informed decisions for irrigation, and implement other agricultural practices. This technology allows farmers to collect data from the field, including soil moisture and weather conditions. Numerous deep learning and machine

* *Corresponding author*: sangeethar.scholar@gmail.com

learning techniques have been developed in the agricultural domain for predictive analytics.

A multihead Long Short-Term Memory (LSTM) model was introduced by Datta *et al.* [1] to successfully forecast soil moisture by minimizing the time required for prediction. The designed model reduces both the mean absolute error and mean square error. However, the space complexity of the prediction remained a major challenging issue. A new encoder-decoder deep learning approach that depends on Long Short-Term Memory (EDT-LSTM) through residual learning was introduced by Li *et al.* [2] for predicting the variation of soil moisture. However, the approach did not demonstrate a significant improvement in accuracy level.

The Multiscale Extrapolative Learning Algorithm (MELA), designed by Chakraborty *et al.* [3], aims to predict soil moisture. However, the dimensionality reduction process was not implemented in this approach, leading to increased complexity in the prediction process. An intelligent system was introduced by Singh *et al*. [4] for precision irrigation, utilizing Machine Learning techniques. However, deep learning was not implemented to enhance prediction accuracy.

A new deep learning model based on LSTM was introduced by Patrizi *et al*. [5] to provide a virtual soil moisture prediction. However, the time required for prediction was not significantly minimized. An LSTM network was introduced by Filipovic et al. [6] for predicting the regional soil moisture with reduced error.

But this approach did not include dimensionality reduction techniques in the analysis of regional soil moisture. An improved spatiotemporal soil moisture modeling approach was developed by Chandrappa *et al.* [7] to enhance the prediction of soil moisture and its inconsistency, aiming to support sustainable irrigation practices. However, this approach did not account for the influence of various environmental conditions on soil moisture prediction. A stacked machine learning algorithm was introduced by Granata *et al.* [8] for detecting multi-step soil moisture with lesser computational time. However, the algorithm did not achieve higher accuracy in predicting soil moisture.

Existing deep learning and machine learning techniques consider several drawbacks such as lower accuracy of soil moisture prediction, higher prediction time, failure to consider error rate, and dimensionality reduction techniques were not carried out. To overcome the problem, the proposed PSLKELC Model is developed to enhance the accuracy and minimize time and error.

The novel contributions or innovation/originality of the proposed PSLKELC Model is described by, the novel PSLKELC Model is introduced to improve the soil moisture prediction from the cloud, based on three major processes namely preprocessing, dimensionality reduction, and classification. The data preprocessing is performed to handle missing and noisy data. Novel Proximity interpolation is applied in data preprocessing to generate new data points for filling in missing values. Noisy data was identified and eradicated by using a novel Tversky coefficient with less time. The novel McNemar statistical test is applied to perform dimensionality reduction to select the most relevant variables with minimum space complexity. To improve the prediction accuracy of the soil moisture, the Laplace kernelized Extreme Learning classifier is used. Testing and training data are examined with a novel Laplace kernel function. The mean absolute error in the prediction process is reduced via the novel Nelder-Mead method. Experiment evaluations are carried out to estimate the performance of our PSLKELC Model and other techniques along with the various metrics.

The paper is structured as follows: Section 2 provides the related works section reviews the related literature and previous research in the field. Section 3 presents a concise description of the PSLKELC Model with an architecture diagram. In section 4, the experimentation process and details of the dataset used are described. The performance results of the proposed technique and existing methods are discussed in 5 using different metrics. Section 6 concludes the paper.

2. Related Works

A fast-learning technique called Extreme Learning Machine (ELM) was developed by Suchithra and Pai [9] for soil moisture prediction. However, the designed model was not efficient in analyzing the soil parameters with higher accuracy. An integration of the linear mixed models and random forests was developed by Makungwe *et al*. [10] for spatial prediction of soil pH. However, the model failed to estimate the parameters of the error. Machine learning (ML) algorithms using IoT devices were introduced by Ikram *et al.* [11] to efficiently and accurately predict crop maximal yield by leveraging IoT devices. However, deep learning-based methods have not been utilized to maintain crop data. The global soil moisture was estimated using machine learning (ML) regression by Jia *et al*. [12]. But, the complexity of the soil moisture estimation was higher.

An attention-based convolutional long short-term memory approach was introduced by Li *et al*. [13] for multistep prediction of soil moisture. Short-term soil moisture prediction was attempted by Dubois *et al.* [14] using a machine learning technique. However, it failed to provide accurate results when tested with additional data. The naive Accumulative Representation model introduced by Basak *et al.* [15] aims to improve the prediction of soil moisture with reduced loss. However, it does not examine the timing and magnitude of responses at different soil moistures over longer time periods.

Sensor based classification and evaluation methods were developed by Florence *et al.* [16] for estimating every churna depending on color, moisture content value, and pH value. But the accuracy was lesser. Deep learning was developed by Wang *et al.* [17] for soil moisture prediction with higher accuracy. Residual-EnDecode-Feedforward Attention Mechanism-LSTM performed by Li *et al.* [18] to obtain water resource utilization. Deep learning regression network (DNNR) was investigated by Cai *et al.* [19] with higher accuracy. A hybrid modeling approach was employed by Zheng *et al.* [20] for forecasting soil moisture content. However, the space complexity was not considered.

3. Methodology

Soil moisture prediction is a significant step in the agriculture domain. Soil moisture prediction refers to the process of forecasting the amount of water content present in the soil. This prediction is crucial in agriculture and environmental sciences since it directly impacts crop growth and irrigation management. By accurately predicting soil moisture levels, farmers can make necessary decisions about irrigation, leading to efficient crop productivity. Internet of Things (IoT) devices are equipped in fields to collect real-time data on soil moisture levels. With the above-said model, PSLKELC is developed to improve soil moisture prediction for precision agriculture.

Fig. 1 given below shows the architecture diagram of the proposed PSLKELC model for accurate soil moisture prediction to assist farmers in crop cultivation. The proposed technique consists of three major processes preprocessing, dimensionality reduction, and classification for improving the accuracy of soil moisture prediction to improve the crop yield. The IoT device is used for collecting the soil parameters and weather conditions soil temperature, soil humidity, soil pressure, soil luminosity, rainfall per day, etc. The collected data are transferred into the cloud for data storage and predictive analytics.

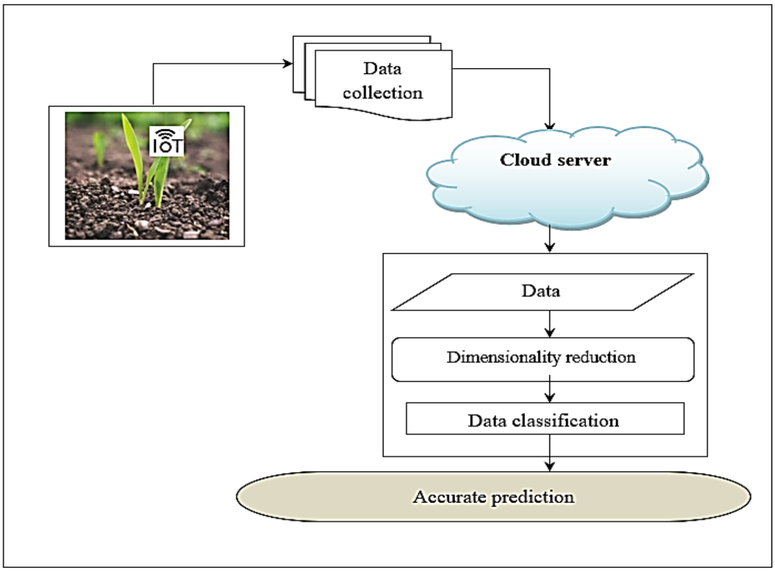

Fig. 1. Architecture of proposed PSLKELC model.

3.1. *Materials and methods*

To perform the soil moisture prediction, the SMART FASAL (Smart Irrigation and Fertilization System for Precision Agriculture using Internet of Things and Cloud Infrastructure) dataset is taken from http://smartfasal.in/ftp-dataset-portal/. This portal stores the real-time soil data and the climate data for three types of crops namely Capsicum, Wheat Dataset, and Rice Dataset. Among them, the rice dataset is considered to perform the soil moisture prediction for precision agriculture using IoT and Cloud. The dataset comprises 13 variables or features and 42666 instances. The aim is to effectively improve soil moisture prediction using the proposed technique.

3.2. *Data preprocessing*

Data preprocessing is a fundamental step in predictive data analysis where raw data is organized and cleaned to make it appropriate for further processing. It involves several essential tasks noisy data removal and missing data handling to ensure data quality and improve the effectiveness of subsequent analysis.

To start with the raw input dataset 'D' and formulated in the form of a matrix as given below.

$$D = \begin{bmatrix} f_1 & f_2 & \dots & f_n \\ X_{11} & X_{12} & \dots & X_{1n} \\ X_{21} & X_{22} & \dots & X_{2n} \\ \vdots & \vdots & \dots & \vdots \\ X_{m1} & X_{m2} & \dots & X_{mn} \end{bmatrix}, m = rows, n = columns \quad (1)$$

From the above input matrix formulation as given in (1), 'n' column features or variables $\{f_1, f_2, \dots f_n\}$ are present with overall sample instances of 'm row respectively, $X_{11}, X_{12} .. X_{mn}$ are data points.

With the above set of features matrix, the first missing data problem is solved by applying a Proximity interpolation. It is a method of finding new data points based on the distance of a discrete set of known data points.

The formula for finding the data points is given below,

$$Q = \frac{|X_a - X_b|}{M_v + 1} \quad (2)$$

Where Q denotes the output of the interpolation method, X_a denotes a data point after the missing data to be filled, X_b denotes a data points before the missing data to be filled, M_v denotes a number of missing value between the two data points, $|X_a - X_b|$ denotes a distance between the data points after the missing data is filled and the data points before the missing data is filled.

The noisy data removal process is performed to minimize the error of the data classification process. Noisy data is meaningless data that is generated due to faulty data collection, data entry errors etc. It is handled by applying a Tversky relationship is a way of measuring how data samples are related or close to each other.

$$T = \frac{[X_i \cap X_j]}{p\,(X_i \cap X_j) + q\,(X_i - X_j)} \quad (3)$$

Where, T indicates a similarity coefficient, X_i and X_j denotes two data points in the particular column, $X_i \cap X_j$ indicates a mutual dependence between the two data points, $X_i - X_j$ indicates a variance between the two data points. From (3), p and q indicates a parameter ($p, q \geq 0$). The coefficient provides the output ranges between [0, 1]. If the coefficient provides an output value of '1,' it denotes a correct data point. On the other hand, if the coefficient provides an output value of '0,' it denotes a noisy data point. These data points are removed to improve the accuracy of the data classification.

3.3. *McNemar statistic multidimensional scaling-based dimensionality reduction*

Dimensionality reduction is a process used in data analytics to reduce the number of variables in a dataset while preserving essential information. McNemar statistic multidimensional scaling is a dimensionality reduction technique by visualizing the level of similarity of individual cases of a dataset using the McNemar statistical test. It is used to measure the pairwise relationship among a set of variables in the dataset.
Let us consider the number of features or variables. $\{f_1, f_2, \dots f_n\}$. The McNemar statistical test is measured between the variables is estimated as follows,

$$\beta = \frac{(f_i - f_j)^2}{f_i + f_j} \tag{4}$$

Where β denotes a McNemar statistical test to measure the relationship between the variables f_i and f_j. The McNemar statistical test provides the output ranges between 0 and 1.

$$Z = \begin{cases} \beta > 0.5\,; select\ the\ variables \\ \beta < 0.5\,; remove\ the\ variables \end{cases} \tag{5}$$

Where Z denotes a multidimensional scaling output, β indicates a statistical test. If the statistical results provide an outcome greater than 0.5, then the variable is selected as relevant. If the statistical results provide an outcome lesser than 0.5, then the variable is selected as irrelevant. In this way, relevant variables are selected for accurate classification, and other features are removed from the dataset. The selected variables are given to the next process resulting in minimizing the time and space complexity of soil moisture prediction.

Algorithm 1, as described above, demonstrates the various processing steps for relevant variable selection based on a similarity measure. Initially, the number of variables is obtained from the dataset. Then, the relationships between variables are measured to determine their relevance. If the similarity coefficient is greater than 0.5, the variable is considered relevant; otherwise, it is considered irrelevant. The relevant variables are then chosen for soil moisture prediction, and the remaining variables are removed from consideration. This process helps to reduce both time and space complexity.

Algorithm 1: McNemar statistic multidimensional scaling-based dimensionality reduction
Input: Pre-processed dataset, number of variables $\{f_1, f_2, \dots f_n\}$ **Output:** Select relevant variables
Begin **1.** Collect the number of variables $f_1, f_2, \dots f_n$ **2.** **For each** variable 'f_i' **3.** **For each** variable 'f_j' **4.** Measure the relationship 'β ' **5.** $\boldsymbol{if}$ $(\beta > 0.5)$**then** **6.** Variable is said to be relevant

7.	**Select** relevant variables
8.	**else**
9.	Variable is said to be irrelevant
10.	**Remove ir**relevant variables
11.	**end if**
12.	**end for**
13.	**End for**
End	

3.4. *Laplace kernelized Extreme Learning classifier-based soil moisture prediction*

Finally, the classification process is performed to achieve accurate soil moisture prediction using the Laplace kernelized Extreme Learning classifier. The proposed classifier is a machine learning algorithm based on a single hidden layer feedforward neural network, offering better performance with an extremely fast learning speed. Unlike conventional ELM, this classifier does not require any iterative training. Instead, the Laplace radial kernel function is applied for data analysis to enhance the classification performance and minimize errors.

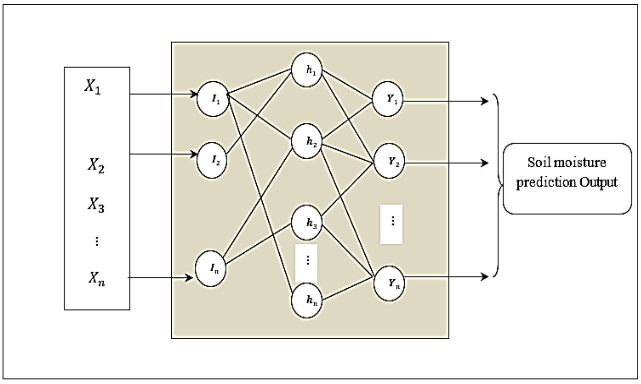

Fig. 2. Construction of Laplace kernelized Extreme Learning classifier.

Fig. 2 illustrates the structure of an Extreme Learning classifier, which is a type of feed-forward neural network used for soil moisture detection through data categorization. The structure consists of an input layer, a hidden layer, and an output layer. The input layer of the Extreme Learning classifier receives the input data (i.e., selected variables) and passes it to the subsequent layers of artificial neurons for further processing. The input layer is located at the beginning of the Extreme Learning classifier. The output of one layer is fully connected to the next successive layer with an equivalent set of weights, forming the entire network.

Let us consider that the training set $\{X, Y\}$ where '$X_1, X_2, \dots, X_n$' indicates training data with the selected variables '$\{f_1, f_2, \dots, f_k\}$' and a label or output 'Y' representing output. The input layer in ELM only receives the input data without performing any computations. On the other hand, the output layer is linear. The important computation process is performed by the hidden layer, which provides results to the output layer.

The weights are fixed and have a straightforward solution that does not require any updating process.

$$I_i = \sum_{i=1}^{n} [X_i(t) * wt_{ij}] + R_b \quad (6)$$

Where, I_i denotes an activity of neurons at the input layer '$X_i(t)$' denotes that the data with weight 'wt_{ij}', and bias function R_b, 'wt_{ij}' denotes a weight between thej^{th} input layer neuron and the i^{th} hidden layer neuron.

The input is transferred into the hidden layer where the training and testing data are analyzed by applying the Laplace RBF kernel. The Laplace RBF kernel '(LK)' is expressed as follows,

$$LK = exp\left(-\frac{\|X_t - X_s\|^2}{D}\right) \quad (7)$$

Where 'D' indicates a deviation, X_t denotes training data which is more similar to the testing data 'X_s' is classified as a particular class. In other words, the computed training data is closer to the testing value being classified as a particular class. The kernel provides the outcomes in the ranges from 0 to 1. Therefore, the hidden layer output is given below.

$$h = \sum_{i=1}^{L} wt_{ij}\ \sigma\,(wt_{jk} h_o + R_b) \quad (8)$$

Where, 'h' represents the result of the hidden layer output, σ indicates an activation function, 'wt_{jk}' denotes the j^{th} hidden layer neuron and k^{th} output layer neuron, wt_{ij} denotes a weight between input and hidden layer, h_o denotes an output of the previous hidden layer, L denotes the number of hidden units.

In order to attain higher accuracy and reduce minimum-error, Nelder–Mead method is applied. It is a numerical method is used to find the minimum of an objective function, such as the error rate in a classification problem.

$$P = arg\,min[act - obsr]^2 \quad (9)$$

Where, P denotes an output of the Nelder–Mead method, arg min denotes an argument of minimum function to minim the error, act denotes actual results and $obsr$ denotes an observed result. Finally, accurate classification results are obtained at the output layer with minimal error. Based on these classification results, soil moisture prediction is performed for agricultural purposes.

Algorithm 2: Laplace kernelized ELM-based soil moisture prediction
Input: selected variables with training data samples, and testing data samples **Output:** Increase the prediction accuracy
Begin **1.** **Number of variables with training data samples** taken at the input layer **2.** **Foreach** training dataX_t**// [hidden layer]** **3.** **Foreach** testing featureX_s **4.** Apply Laplace kernel function 'LK' **5.** **end for** **6.** **end for** **7.** **Classify** data samples into particular class **8.** Apply Nelder–Mead method to minimize the error rate using (9) **9.** Obtain the classification results at the output layer **End**

Algorithm 2 describes the process of soil moisture prediction using a Laplace kernelized ELM, which aims to achieve higher accuracy and minimize time consumption. The proposed learning classifier involves multiple layers to analyze the soil data. The training data samples are fed into the input layer, where random weights and biases are assigned. Next, the hidden layer conducts data analysis using the Laplace Radial Basis Function (RBF) kernel. Based on the kernel output, the classification is performed. Finally, the Nelder-Mead method is applied to minimize the error and obtain the final prediction results at the output layer.

4. Experimental Evaluation

Experimental assessments of the PSLKELC and existing multihead LSTM model [1] and EDT-LSTM [2] are implemented in Python with SMART FASAL (Smart Irrigation and Fertilization System for Precision Agriculture using Internet of Things and Cloud Infrastructure) dataset. The dataset collected from http://smartfasal.in/ftp-dataset-portal/. The soil information about the rice dataset is considered to conduct the experiment. The number of features or variables description is listed in Table 1.

Table.1. Feature description.

S. No	Feature	Description
1	Sensor ID	-
2	Soil_moisture 1	Acquires information from the sensors installed within the soil at a depth level 15cms
3	Soil_moisture 2	Acquires information from the sensors installed within the soil at a depth level 45cms
4	Soil_moisture 3	Acquires information from the sensors installed within the soil at a depth level 80cms
5	TEMP	Soil temperature
6	HUMD	Soil humidity
7	PRSR	Soil pressure
8	LMNS	Soil Luminosity
9	Rainfall	Rainfall per day (mm)
10	week cycle count	Week cycle count of recording
11	Day	day of recording
12	Date	Date of recording (DD: MM;YY)
13	Time IST	Time of recording

5. Comparative Analysis

In this section, the comparative analysis of the proposed PSLKELC and existing multihead LSTM model [1] and EDT-LSTM [2] are discussed with different evaluation metrics such as accuracy, mean absolute error, time complexity, and space complexity.

5.1. *Metrics to evaluate the results*

Prediction Accuracy: It is defined as the ratio of the number of data samples that are correctly predicted to the total number of data samples. Accuracy is formulated as given below,

$$PR\_ACC = \sum_{i=1}^{n} \left[\frac{CCS}{X_i}\right] * 100 \quad (10)$$

Where 'PR_ACC' denotes the prediction accuracy, 'CCS' represents the number of data samples correctly predicted and 'X_i' indicates the total number of data samples. The accuracy is measured in percentage (%).

Mean absolute error: It is a metric used to measure the error of a predictive model. It quantifies the average difference between the predicted values and the actual values. Mean absolute error is formulated as follows:

$$ER_{MA} = \frac{1}{n}\left|y_{act} - y_{pred}\right| \quad (11)$$

Where 'ER_{MA}denotes the mean absolute error, 'y_{act}represents the number of actual results, y_{pred} indicates a predicted value

Time complexity: It is formulated as the time taken to accurately predict soil moisture through data sample classification. The time is calculated as follows:

$$T_{Comp} = \sum_{i=1}^{n} X_i * t(P) \quad (12)$$

Where 'T_{Comp}' indicates the time complexity, X_i represents the number of data samples '$t(P)$]' denotes the time for predicting the single data samples. Therefore, the time is measured in terms of milliseconds (ms).

Space complexity: It is formulated as the amount of memory space required to accurately predict soil moisture through data sample classification. The calculation of space complexity is as follows:

$$SP_{Comp} = \sum_{i=1}^{n} X_i * Mem(P) \quad (13)$$

Where 'SP_{Comp}' indicates the space complexity, X_i represents the number of data samples '$Mem(P)$' denotes the memory for predicting the single data samples. Therefore, the time is measured in terms of Mega Bytes (MB).

Table 2. Comparing existing and proposed system values.

Metrics	Methods		
	PSLKELC	Multihead LSTM model	EDT-LSTM
Prediction accuracy (%)	96.5	94.2	92.7
Mean absolute error	0.035	0.058	0.073
Time complexity (ms)	115	128	153
Space complexity (MB)	135	165	185

Table 2 presented above illustrates the comparative performance analysis of soil moisture prediction accuracy concerning the number of data samples taken from the dataset. The results reveal different performance metrics for all three methods, namely PSLKELC, the existing multihead LSTM model [1], and EDT-LSTM [2]. The observed values indicate that the PSLKELC technique outperforms the other two existing methods for all the metrics evaluated. The results are illustrated in the graphical analysis.

Fig. 3 depicts a graphical comparison analysis of prediction accuracy for three methods: PSLKELC, the existing multihead LSTM model [1], and EDT-LSTM [2]. The x-axis represents the methods, and the y-axis represents the accuracy of soil moisture prediction. The graph shows a downward trend, indicating that the proposed method has higher accuracy than the existing models [1,2]. This improvement is achieved in the PSLKELC model, which utilizes a Laplace kernelized Extreme Learning classifier. The learning classifier uses the Laplace kernel function to perform data analysis in a cloud environment with the help of training data samples and testing data samples, enabling accurate predictions. As a result, the overall prediction accuracy of the PSLKELC model is increased by 2 % and 4 % when compared to models [1,2], respectively.

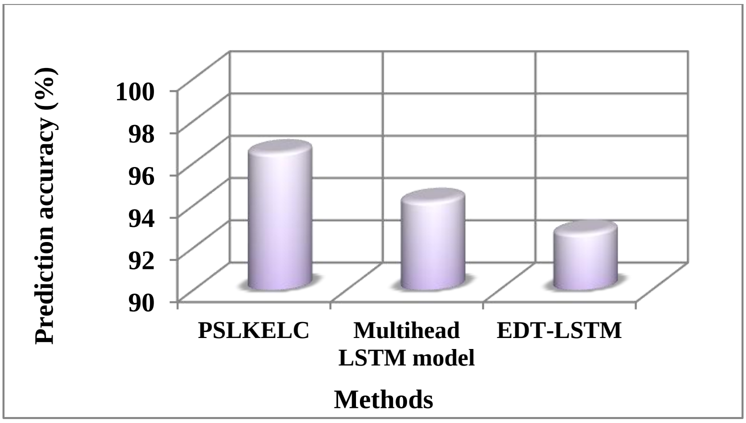

Fig. 3. Comparison of prediction accuracy.

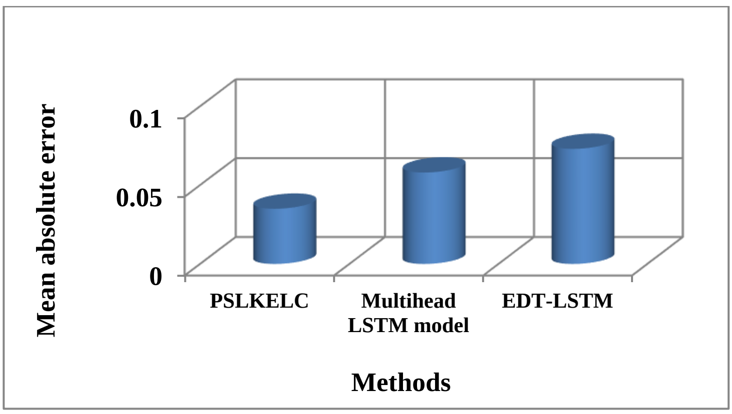

Fig. 4. Comparison of mean absolute error.

Fig. 4 presents a performance analysis of mean absolute error using three different methods: PSLKELC, the existing multihead LSTM model [1], and EDT-LSTM [2]. As depicted in Figure 4, the PSLKELC model shows significantly minimized mean absolute error compared to the existing methods. This improvement is achieved through the utilization of the Laplace kernelized Extreme Learning classifier, which employs the Nelder-Mead method to minimize the error between actual and predicted classification results. Finally, the overall comparison results demonstrate that the mean absolute error performance of the PSLKELC is remarkably reduced by 39 % and 52 % when compared to the existing methods [1,2], respectively.

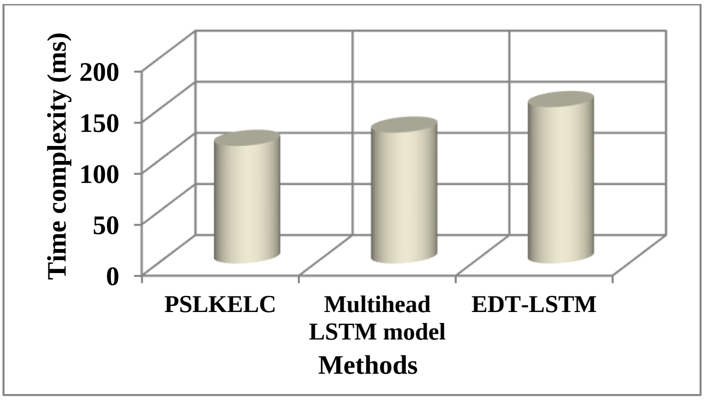

Fig. 5. Comparison of time complexity.

Fig. 5 illustrates the performance outcomes of time complexity involved in soil moisture prediction in the cloud. The graphical results demonstrate that the proposed PSLKELC model exhibits superior performance compared to conventional methods. As depicted in Fig. 5, the horizontal axis represents the methods used in the experimental process, while the vertical axis displays the time complexity of soil moisture prediction using three methods: PSLKELC model, multihead LSTM model [1], and EDT-LSTM [2]. This improvement is achieved through the data preprocessing and dimensionality reduction process of the PSLKELC model. Initially, data preprocessing is conducted to handle missing values and noisy data in the given dataset. Proximity imputation is applied to address the missing values, while a Tversky coefficient is used to measure the similarity between data points in the dataset. Noisy data is then removed based on the similarity value. Additionally, McNemar statistic multidimensional scaling-based dimensionality reduction is performed to select significant variables from the dataset for soil moisture prediction.

Figure 6 presents a performance comparison of space complexity among different methods, including PSLKELC, the existing multihead LSTM model [1], and EDT-LSTM [2]. The horizontal axis represents the three different methods, while the vertical axis illustrates their respective space complexities. Especially, the proposed PSLKELC model demonstrates superior performance compared to the existing methods [1,2]. The PSLKELC model achieved this by utilizing the McNemar statistic multidimensional scaling-based dimensionality reduction technique, which effectively minimizes the space complexity for soil moisture prediction. The McNemar statistical test is employed to measure the relationship between variables in the given dataset. The output of the test result greater than 0.5 are considered relevant, while those with a result below 0.5 are considered irrelevant. The relevant variables are utilized for soil moisture prediction, and the remaining variables are removed, contributing to the overall reduction in space complexity.

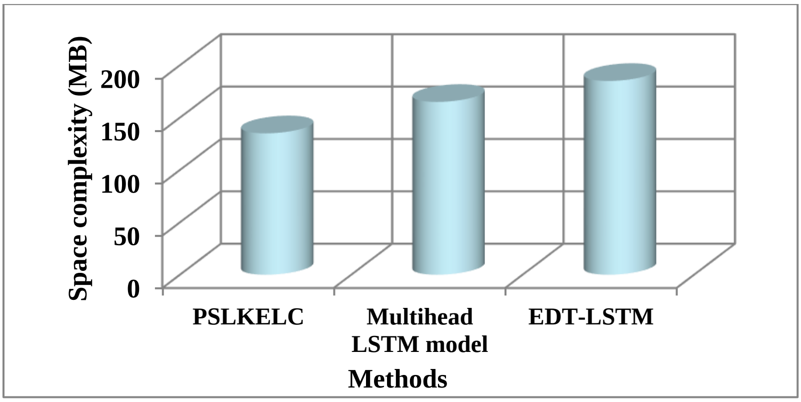

Fig. 6. Comparison of space complexity.

6. Conclusion

The manuscript has presented a novel IoT-aware PSLKELC model for soil moisture prediction in the cloud. The objective of the proposed PSLKELC model is to accurate longer-term soil moisture predictions for improving crop yield with less time. The proposed PSLKELC model is designed with preprocessing, dimensionality reduction, and classification. IoT device is employed to gather the soil parameters and weather conditions soil temperature, soil humidity, soil pressure, soil luminosity, rainfall per day, etc. The gathered data are sent into the cloud for data storage and predictive analytics. At first, the data preprocessing step is to clean and structure the dataset. Next, the most significant variables are chosen to perform the dimensionality reduction process for soil moisture prediction. Lastly, the data samples are classified by Laplace kernelized Extreme Learning Classifier for soil moisture prediction. Laplace kernel function is employed for examining testing and training data. Error is decreased via the Nelder–Mead method. In this way, the accurate prediction is achieved. The performance result of the proposed PSLKELC model and two existing methods are carried out using Python with the aid of the SMART FASAL dataset. The proposed PSLKELC model is compared with two existing algorithms namely the multihead LSTM model and EDT-LSTM based on numerous data samples. The simulation consequences validated that the PSLKELC model provides better results in performance metrics like accuracy, time, error, and space complexity compared to conventional methods. From the estimated result, the proposed PSLKELC model is to provide better performance of higher prediction accuracy by 3%, and minimal error by 46% than compared to existing methods. Also, the time and space complexity are minimized when compared to state-of-art-works. The proposed PSLKELC model is based on a few parameters, is very accurate, and has very limited precision and recall. In the future, the proposed model is further extended to consider the precision and recall metrics for soil moisture prediction.

Acknowledgment

The authors are grateful to the management and Principal for their motivation and constant support. This research has been supported by the grant obtained under the scheme of Seed Money for Research Projects from Cauvery College for Women (Autonomous), Tiruchirappalli – 620 018, India.

References

1. P. Datta and S. A. Faroughi, Geoderma **433**, ID 116453 (2023). https://doi.org/10.1016/j.geoderma.2023.116452
2. Q. Li, Z. Li, W. Shangguan, X. Wang, L. Li, and F. Yu, Comput. Electr. Agricult. **195**, ID 106816 (2022). https://doi.org/10.1016/j.compag.2022.106816
3. D. Chakraborty, H. Başagaoglu, S. Alian, A. Mirchi, D. N. Moriasi, P. J. Starks, and J. A. Verser, Expert Syst. Applicat. **213**, ID 119056 (2023). https://doi.org/10.1016/j.eswa.2022.119056
4. D. K. Singh, R. Sobti, P. K. Malik, S. Shrestha, P. K Singh, and K. Z. Ghafoor, Security Commun. Networks **2022**, ID 7283975 (2022). https://doi.org/10.1155/2022/7283975
5. G. Patrizi, A. Bartolini, L. Ciani, V. Gallo, P. Sommella, and M.Carratu, IEEE Transact. Instrument. Measurement **71**, ID 2515411 (2022). https://doi.org/10.1109/TIM.2022.3196446
6. N. Filipovic, S. Brdar, G. Mimic, O. Marko, and V. Crnojevic, Biosyst. Eng. **213**, 30 (2022). https://doi.org/10.1016/j.biosystemseng.2021.11.019
7. V. Y Chandrappa, B. Ray, N. Ashwatha, and P. Shrestha, Internet Things **21**, ID 100671 (2023). https://doi.org/10.1016/j.iot.2022.100671
8. F. Granata, F. D. Nunno, M. Najafzadeh, and I. Demir, Hydrology **10**, 1 (2022). https://doi.org/10.3390/hydrology10010001
9. M. S. Suchithra and M. L. Pai, Inform. Process. Agricult. **7**, 72 (2020). https://doi.org/10.1016/j.inpa.2019.05.003
10. M. Makungwe, L. M. Chabala, B. H. Chishala, and R. M. Lark, Geoderma **397**, ID 115079 (2021). https://doi.org/10.1016/j.geoderma.2021.115079
11. A. Ikram, W. Aslam, R. H. H. Aziz, F. Noor, G. A. Mallah, S. Ikram, M. S. Ahmad, A. M. Abdullah, and I. Ullah, J. Sensors **2022**, 1 (2022). https://doi.org/10.1155/2022/2022923
12. Y. Jia, S. Jin, H. Chen, Q. Yan, P. Savi, Y. Jin, and Y. Yuan, IEEE J. Selected Topics Appl. Earth Observat. Remote Sensing **14**, 4879 (2021). 10.1109/JSTARS.2021.3076470
13. L. Li, Y. Dai, W. Shangguan, N. Wei, Z. Wei, and S. Gupta, J. Hydrometeorol. **23**, 337 (2022). https://doi.org/10.1175/JHM-D-21-0131.1
14. A. Dubois, F. Teytaud, and S. Verel, Comput. Electron. Agricult. **180**, ID 105902 (2021). https://doi.org/10.1016/j.compag.2020.105902
15. A. Basak, K. M. Schmidt, and O. J. Mengshoel, Int. J. Data Sci. Anal. **15**, 9 (2023). https://doi.org/10.1007/s41060-022-00347-8

16. J. R. Florence, S. S. Priyadharsini1, G. S. Chandran, J. Sci. Res. **14**, 1 (2022). 10.3329/jsr.v14i1.54739
17. Y. Wang, L. Shi, Y. Hu, X. Hu, W. Song, and L. Wang, Hydrol. Earth Syst. Sci. **28**, 4 (2023). https://doi.org/10.5194/hess-2023-177
18. X. Li, X. Li, Z. Zhang, Q. Li, and J. Zhu, Water **16**, 10 (2023). https://doi.org/10.3390/w16101376
19. Y. Cai, W. Zheng, X. Zhang, L. Zhangzhong, and X. Xue, PloS One **14**, 4, (2019). https://doi.org/10.1371/journal.pone.0214508
20. W. Zheng, K. Zheng, L. Gao, L. Zhangzhong, R. Lan, L. Xu and J. Yu, Agronomy **14**, 3, (2024). https://doi.org/10.3390/agronomy14030432