Journal of Information Systems Engineering and Management
2025, 10(44s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Enhancing River Water Quality Prediction: A Hybrid Approach of Deep Autoregression and Bidirectional LSTM Networks

Dr. D. Narmadha[1], Dr. Mythily M[2], Dr. G. Gifta Jerith[3], Dr. Sweety Julia[4], Dr. G. Naveen Sundar[5*], Dr. S. Joyson Abraham[6]

[1]*Assistant Professor, Division of AIML, Karunya Institute of Technology & Sciences*
narmadha@karunya.edu
[2]*Associate Professor, Division of CSE, Karunya Institute of Technology & Sciences*
mythily@karunya.edu
[3]*Assistant Professor, Department of AI&ML, School of Engineering, Malla Reddy University, Hyderabad*
ggiftajerith@gmail.com
[4]*Assistant Professor, Department of AI&ML, School of Engineering, Malla Reddy University, Hyderabad*
sweetyjulia2221@gmail.com
[5*]*Associate Professor, Division of CSE, Karunya Institute of Technology & Sciences*
naveensundar@karunya.edu
[6]*Associate professor, Department of Aeronautical engineering, Malla Reddy College of Engineering and technology, Hyderabad*
joysonabraham19@gmail.com

ARTICLE INFO

Received: 29 Dec 2024

Revised: 12 Feb 2025

Accepted: 27 Feb 2025

ABSTRACT

Time series forecasting models are essential for the prediction of river water quality, which is an important part of maintaining environmental sustainability and protecting public health. However, traditional prediction methods often fail to capture the complex temporal dependencies and non-linear relationships that are inherent to river water quality information. An innovative approach is proposed here, to improve river water quality prediction by combining deep autoregression and bidirectional LSTM networks. The deep autoregression model takes advantage of the inherent temporal correlations that are present in the river water quality data and integrates autoregressive features that encapsulate historical information about the target variable. The BiLSTM architecture lets the model to understand from the historical and future contexts, allowing a holistic view of the temporal dynamics that affect water quality in the river basin. We also introduce feature augmentation technique that improves the proposed model's ability to capture the complex relationships and patterns in the data. By enriching the model's input features with domain specific information, including meteorological data and hydrological parameters as well as land use characteristics that are relevant to the river basin. By incorporating domain-specific information through feature augmentation, the model is poised to provide more accurate and insightful predictions, thereby aiding in effective water resource management and environmental conservation efforts. The error rate has been approximately reduced by 42.6% by using the DAFA-BiLSTM model, when compared to the conventional LSTM models.

Keywords: Hybrid Approach, Autoregression, LSTM

I. INTRODUCTION AND RELATED WORKS

The modern world places a great deal of importance on water quality. Both human health and the ecosystem are severely impacted. The demand for high-quality water has increased due to the growth in the global population. The rate at which groundwater resources are being depleted by humans has increased due to increased demand. Poor management of water resources, an increase in industries, the disposal of unwanted materials in water bodies, and many other factors are the causes of the declining water quality. The majority of diseases have water contamination as one of their main causes. Numerous hazards to life are brought about by inadequate water management and sanitary facilities. Because of this, a system that can precisely estimate the water quality is required so that it can be

used for a variety of applications. The quality is affected by changes over time in different ways, which makes it challenging for the models that are currently in use to predict accuracy. These models have intricate parameter requirements, and any deviation from these requirements would result in significant adjustments to the accuracy and error rate. It has been proposed that this study will enhance the Burnett River's water quality.
Seasonal fluctuations and land use are two of the many elements that have a big influence on water quality. An attention-based long short-term memory (LSTM) model for water quality prediction was presented by Honglei Chen et al. [1]. For this, information was acquired about the locations of automated monitoring stations and the limits of catchments. In order to improve prediction accuracy, a comparison between LSTM and AT-LSTM findings was done for the model, which attempted to estimate river dissolved oxygen level in both single and multiple stages. The F-score was used to assess the anticipated groundwater index quality using an LSTM RNN model, and the results showed a good agreement with the actual values [2]. [3] A different LSTM model that makes use of deep learning was put forth to forecast pH and water temperature using information from a mariculture base located in Xincun Town, China.

To resolve this issue, an AT-BiLSTM technique was employed. The attention layer and bidirectional feature extraction employed by the model significantly affect the predicted values of the water quality data. The time series' water quality variables were forecasted using LSTM, CNN, transfer learning, and attention networks [5]. In comparison to other machine learning models, the accuracy prediction is greatly influenced by the variables and the transformation method applied to the dataset. Several techniques, including SVR, ARIMA, LR, and backpropagation, are used to predict the linear relations between the water quality data. It is believed that the data is stable. Only the linear relationship between the data is captured by these models. Jing Bi et al. [6] suggested a hybrid approach to prediction. Two real-world datasets are used to assess how accurate this model predicts. The best prediction results are produced by this model, which addresses the non-linear characteristics of the data. Utilising a model known as Graph Convolutional Network with Feature and Temporal Attention [7], the multivariate water quality data was predicted. Dissolved oxygen data is the type of data used. [8] This model takes care of the data's temporal dependencies. It is employed to uncover the hidden correlations between the indicators of water quality. To forecast the lake's dissolved oxygen content, an ensemble model was created. This model supported agriculture and assisted in regulating the dissolved oxygen content. The DO prediction result is obtained by adding the predicted values of each component. The prediction was highly accurate and controlled the data's temperature more effectively [9].

[10] A model based on IGRA and LSTM was introduced considering the information on the multivariate correlation. This model was applied to two water quality datasets, which proved that the method can provide better predictions compared to the other prediction methods, even with the multivariate correlations and time sequence of the information available. [11] A neural network model, with CNN and LSTM, is introduced, to estimate the dissolved oxygen index. The model extracted the local features of the data and chose optimal parameters to train the model. The results were more accurate than the conventional models proposed earlier. A CNN-LSTM hybrid model was proposed for the prediction of the short-term power load. The features are extracted using CNN. The feature vectors were constructed based on the time series [12]. This was used as input for the LSTM model. The quality of water was determined using traditional machine learning approaches such as Adaptive Boosting (Ada-Boost), and SVC, RF (Random Forest XGB (XGBoost), DTC (Decision Tree). The work was done based on water quality parameters such as pH, Turbidity, Solids, Sulfate, Organic carbon, Hardness, Trihalomethanes, and Conductivity. These parameters were tested according to the WHO standards [13]. Ye et al. introduced a model which can predict the pollutant index with the data obtained from the rivers in Shanghai [14]. A short-term photovoltaic power forecasting system based on an LSTM and attention mechanism was presented in order to predict short-term solar power generation in a sequential form [15].

Because of their better forecasting performance over statistical techniques, support vector regression (SVR), dual-orthogonal radial basis function network, and Elliott wave pattern-based artificial neural network (ANN) are becoming more and more well-liked in the prediction community. They are very good at catching features, generalisation, and self-learning. Unfortunately, due to their lack of memory, dynamic capabilities, and inability to take time series correlations into account, these machine learning techniques—along with feed-forward ANN (FFANN)—frequently perform poorly in time series prediction tasks. For sequential systems with memory, recurrent

Journal of Information Systems Engineering and Management
2025, 10(44s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

neural networks (RNNs) are recommended over FFANNs. Even though RNNs are among of the finest models for time series forecasting, their capacity to predict sequences is limited by their inability to handle chaotic, high-dimensional, and long-range dependent time series datasets.

A new method is presented to improve feature representations in the BiLSTM network and increase prediction performance for time series datasets: the deep autoregression feature augmented bidirectional LSTM network (DAFA-BiLSTM). The VA layer in the DAFA-BiLSTM architecture is used to extract both linear and nonlinear feature observations of the input sequence. The top layer of a deep augmented BiLSTM (DA-BiLSTM) module processes the nonlinear combination feature vectors from the VA layer after receiving these observations. Time-delayed linear feature vectors are incorporated as additional input signals into each BiLSTM layer, while nonlinear combination feature vectors are immediately received by the bottom layer from the VA layer. The nonlinear combinations that are used as additional inputs at each layer are easily handled by the deep BiLSTM. Linear feature vectors are added directly as input to the deep BiLSTM in order to preserve all features. The DAFABiLSTM consists of several DA-BiLSTM layers as well as a VA module, in contrast to earlier deep LSTM models. The dynamics of every nonlinear observation that the VA module retrieves are captured by each BiLSTM layer. In the DA-BiLSTM architecture, the output of each BiLSTM layer is concatenated with external time-delayed linear input vectors from the VA layer to provide new extra inputs for successive neighbouring layers.

II. PROPOSED SYSTEM

1. *BiLSTM*

The gradient issue with conventional RNNs—where gradients would either disappear or get incredibly tiny during training—was the reason LSTMs were developed. In order to address this problem, long-term dependencies present in time series data were captured using LSTM cells. The neural network's hidden layers are these cells. LSTM cells are good in capturing long-term dependencies, in contrast to conventional RNNs. They are made up of forget, output, and input gates, which control the information flow that is essential for output prediction. LSTM preserves the current states of recurrent neurons based on both prior neuron states and present inputs, while different components alter the structure of recurrent units.

The forget gate decides whether information should be remembered or forgotten based on input from the previous cell output and the current input. By seeing patterns in the input vectors that need to be included, the input gate is essential in determining the shape of the cell state. In the meantime, pertinent data is chosen by the output gate to be output from the current cell state. The following are the formulas that explain the LSTM operations:

$$I_t = \sigma\,(w_i x_t + W_i h_{t-1} + b_i) \qquad (1)$$

$$F_t = \sigma\,(w_f x_t + W_f h_{t-1} + b_f) \qquad (2)$$

$$O_t = \sigma\,(w_o x_t + W_o h_{t-1} + b_o) \qquad (3)$$

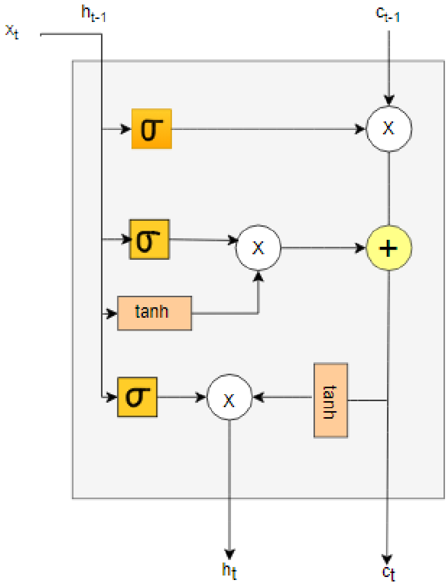
$$\acute{C}_t = tanh\,(w_c x_t + W_c h_{t-1} + b_c) \qquad (4)$$

$$C_t = F_t \otimes C_{t-1} + I_t \otimes \acute{C}_t \qquad (5)$$

$$H_t = O_t \otimes tanh\,(C_t) \qquad (6)$$

At time step t, the sequential input and recurrent output states are denoted using x_t and h_t, respectively. In the above equations, w_i, w_f, w_o and w_c are used to denote the recurrent weight matrices of the gates. b_i, b_f, b_o and b_c are used to represent the biases. $\acute{C}_t$ is used to represent the state of the candidate cell, which is utilised to update the initial memory cell C_t. The hyperbolic tangent function is represented by *tanh*, and the activation function is represented by σ.

Journal of Information Systems Engineering and Management
2025, 10(44s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article



During the training process, the LSTM might lose feature information as it only takes input vectors in a single direction into account. This makes it difficult for the sequence information to be thoroughly analyzed. To solve this BiLSTM is designed to record the data in both forward and backward directions. BiLSTM consists of two LSTMs that are present parallel to each other, wherein one LSTM processes the input forward while another processes it backward direction. Each LSTM has its own hidden state, which is linked in a sequential manner to provide the output. The one that processes the input in the forward direction captures the dependencies from past to future, whereas, the other LSTM captures the dependencies from future to past. This allows us to understand the input sequence in an efficient manner in both past and future context.
The forward direction LSTM, stores the information in $h_{f(t)}$ using the past time series values; the backward direction LSTM, store the information in $h_{b(t)}$ using the future sequence values. These states are linked in a sequential manner to produce the final output. The equations of these states are given below:

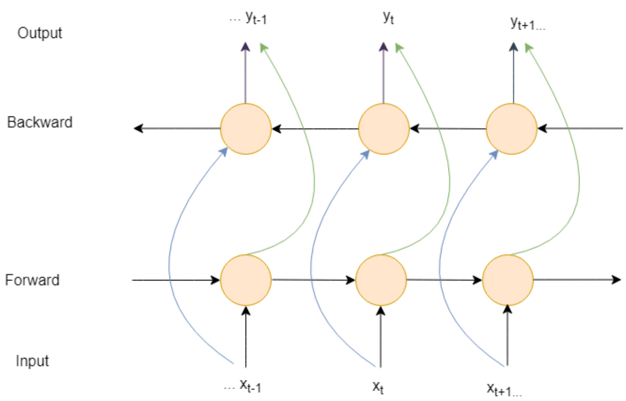
$$h_{f(t)} = tanh\ (w_{fh}x_t + w_{fhh}\ h_{f(t-1)} + b_f) \quad (7)$$

$$h_{b(t)} = tanh\ (w_{bh}x_t + w_{bhh}\ h_{b(t-1)} + b_b) \quad (8)$$

The forward and backward weights are represented using weight matrices w_{fh} and w_{bh}. These represent the weights from the recurrent unit to itself. The bias signals of both directions are represented by b_f and b_b. The recurrent layer activation function that is used here is *tanh*. The output is generated as a vector, which is given below:

$$Y_t = \sigma\ (w_{fhy}h_{f(t)} + w_{bhy}h_{b(t)} + b_y) \quad (9)$$

w_{fhy} and w_{bhy} are used to represent the forward and backward weights from the internal unit to the output. The output layer uses the sigmoid activation function which is denoted by σ, and the bias is represented using b_y.



Journal of Information Systems Engineering and Management
2025, 10(44s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

2. *Vector autoregression*

An statistical model known as vector autoregression is a multivariate version of the univariate autoregressive model. Multiple time series variables, the autoregressive process, the model's order, and the matrix formulation are the essential elements of a vector autoregression model. It depicts the time series variables' dynamic interdependencies, which are crucial for comprehending complex systems. Vector autoregression is a viable solution for the problem of nonlinear pattern recognition. Vector autoregression works on the basis of processing input signal in a high dimensional space of features, thereby avoiding solving the mapping in a high dimensional space of features.

3. *Deep autoregression feature augmented BiLSTM*

Deep Recurrent Neural Networks (RNNs) are an efficient method because they can extract features from datasets of time series. The DAFA-BiLSTM, which is suggested in this research, is made up of several BiLSTM layers joined by an augmented structure. It combines the benefits of both BiLSTM networks and deep autoregression. Time series forecasting jobs benefit greatly from the use of external features and past sequential patterns. The suggested system consists of two main components: a vector autoregression component and a deep enhanced BiLSTM-based prediction model. Lagged values of the target variable and external features make up the input layer. To extract the intricate temporal correlations from the lagged target data, the autoregression component makes use of several layers. While the output layer produces the predicted values based on the learnt representations, the BiLSTM layer records the sequential patterns.
Deep BiLSTM models are able to learn more effectively thanks to vector autoregression, which offers an efficient feature representation technique. Additionally, this enhances the suggested model's performance with time series datasets. This approach improves the proposed system's interpretability and makes the system's information flow directly known. Complex temporal dependencies in the dataset can be effectively captured by combining deep autoregression with BiLSTM. Multiple BiLSTM layers are part of the increased structure of the proposed system. Vector autoregression was used to change the feature vectors, and each layer can dynamically convert those vectors into features. Time-lagged linear features are added to each BiLSTM layer and the output of each BiLSTM layer is gradually learned to influence the outcomes of the proposed system. In comparison to the old models, this feature increases the flexibility and predictability of the proposed system.

4. *Working of the proposed system*

The DAFA-BiLSTM model is the outcome of improving the suggested design by combining several BiLSTM layers into an upgraded structure. Comparing this model to traditional BiLSTM models reveals a number of benefits. First off, the attention method allows the model to dynamically concentrate on pertinent portions of the input sequence when making predictions. With this method, the model is better able to identify long-term connections and subtle patterns in the data, which is important for tasks like predicting pH levels in time-series forecasting applications. Second, the DAFA-BiLSTM model uses a technique called dropout regularisation, which randomly disconnects connections between neurons during training to assist prevent overfitting.
This regularisation strategy lessens the model's propensity to memorise noise found in the training set, which helps to improve the model's generalisation abilities. Last but not least, the DAFA-BiLSTM model efficiently learns from both past and future contexts concurrently by capturing bidirectional relationships within input sequences by utilising both forward and backward LSTM layers. The DAFA-BiLSTM model performs better than typical BiLSTM models in part because of its design, especially when it comes to tasks like time-series prediction that require the study of sequential data. The architecture of the DAFA-BiLSTM model, which is suggested to improve the BiLSTM model's performance, is shown in Figure 3.4.

Journal of Information Systems Engineering and Management
2025, 10(44s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

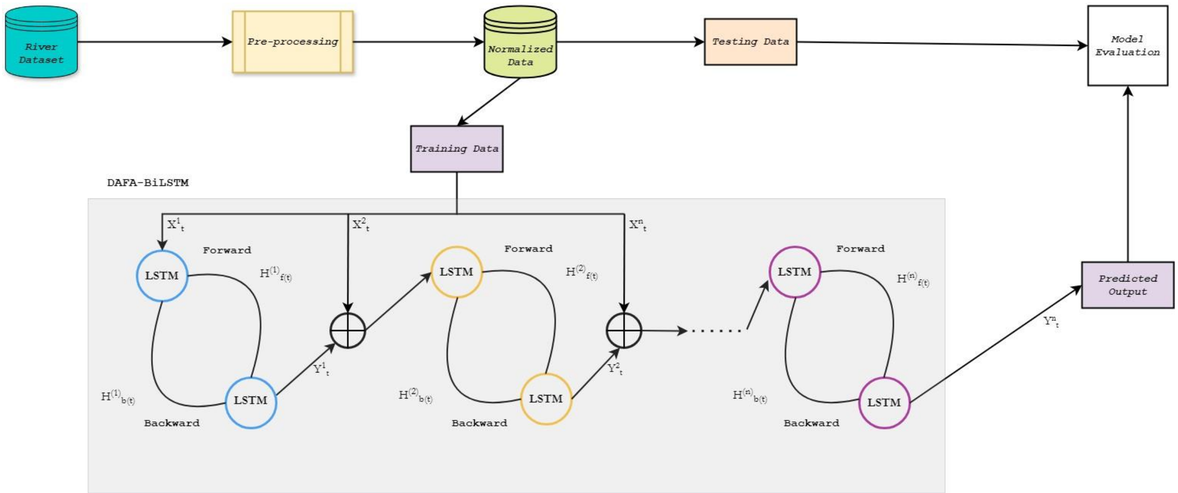


Figure 3.4: Proposed Architecture of DAFA-BiLSTM

The Deep Autoregression Feature Augmentation BiLSTM (DAFA-BiLSTM) model integrates several advanced techniques to enhance its predictive capabilities, particularly in time-series forecasting tasks. The Bidirectional Long Short-Term Memory (BiLSTM) architecture, which is skilled at capturing temporal dependencies within sequential data, is the fundamental component of the model. Unlike traditional LSTM models that only consider past information, BiLSTM incorporates both past and future context, enabling more comprehensive learning. However, what sets DAFA-BiLSTM apart is its incorporation of an attention mechanism and deep autoregression feature augmentation. The attention mechanism helps the model focus more on pertinent data when making predictions by enabling it to dynamically determine the relative importance of various input sequence segments. Long-range dependencies and subtle patterns in the data are better captured by the model thanks to this attention mechanism, which can be important for precise forecasting. The deep autoregression feature augmentation technique further enriches the model's representation by augmenting the input features with their autoregressive counterparts. This augmentation enables the model to explicitly capture the temporal relationships between consecutive time steps, empowering it to better understand the underlying dynamics of the time series data. By incorporating both deep autoregression and attention mechanisms into the BiLSTM architecture, DAFA-BiLSTM leverages the strengths of each component to achieve superior predictive performance compared to traditional BiLSTM models. By combining the strengths of BiLSTM, attention mechanisms, and deep autoregression feature augmentation, DAFA-BiLSTM offers a robust framework for accurate and reliable forecasting of the water quality.

The dataset is normalized using min-max normalization. It is also called as feature scaling, a preprocessing technique that is used to scale the data to a particular range. It prevents the features that have larger sizes from affecting the model training. For a particular feature xi, with the range of [a,b] the min-max normalization is represented in Equation 2:

$$x_i' = \frac{x_i - \min(X)}{\max(X) - \min(X)}(b - a) + a \qquad (2)$$

Where x_i' represents the normalized value, xi is the original value, min(X) represents the minimum value and max(X) represents the maximum value of the feature, a and b represent the desired range for the values.

4.1.Performance Metrics

The metrics used to evaluate the model prediction are the NMSE, R2 score and MAPE values.

Journal of Information Systems Engineering and Management
2025, 10(44s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

4.1.1. NMSE

It is a variation of MSE (Mean Square Error). It normalizes the error by dividing it by the variance of the actual values. This make it an unitless measure, which allows it to be used for different datasets. The formula for NMSE is:

$$NMSE = \frac{\sum_{i=1}^{n}(y_i - \hat{y}i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{3}$$

4.1.2. R2 score

It is the measure that indicates how well the actual value is predicted by the model. The formula for R2 score is:

$$R2 = 1 - \frac{\sum_{i=1}^{n}(yi - \hat{y}_1)^2}{\sum_{i=1}^{n}(y_; - \hat{y}i)^2} \tag{4}$$

4.1.3. MAPE

It is the mean percentage difference between the actual and expected values. The formula for MAPE is:

$$MAPE = \frac{1}{n}\sum_{1=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_1}\right| 100 \tag{5}$$

4.1.4. RMSE

It measures the average magnitude of the errors between predicted values and actual values. RMSE is particularly useful because it penalizes large errors more heavily than smaller ones, making it sensitive to outliers. The formula for RMSE is:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{6}$$

III. RESULTS AND DISCUSSIONS

The application of the LSTM network has demonstrated better results compared to traditional machine learning techniques. LSTMs are the type of recurrent neural networks that are better at capturing temporal dependencies and are the best RNN which can be used for time-series data. This makes it effective for modeling and predicting the parameters over time. It learns complex patterns and relationships within the given data and enables us to learn and predict them. The main aim of the study was to find out which model provides better values, the conventional LSTM models or the proposed DAFA-BiLSTM model, to predict the quality of water. Figure 4 shows the comparison between the actual and expected values of the LSTM model. The conventional LSTM models have lower performance compared to that of the DAFA-BiLSTM model.

Table 2. NMSE, RMSE, R2 and MAPE scores

Model	NMSE	RMSE	R2	MAPE
LSTM	0.309	0.384	0.69	3.56
BiLSTM	0.297	0.377	0.70	3.38
DAFA-BiLSTM	0.261	0.216	0.84	1.74

Journal of Information Systems Engineering and Management
2025, 10(44s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

Table 2 represents the NMSE, RMSE, MAPE and R2 scores of different models which is used to compare the results and prove which model has the better performance. DAFA-BiLSTM appears to be more effective when compared to the other models, with the lower error rate. Figure 4 and 5 are similar to Figure 3, but they show the values of the BiLSTM and DAFA-BiLSTM model respectively.

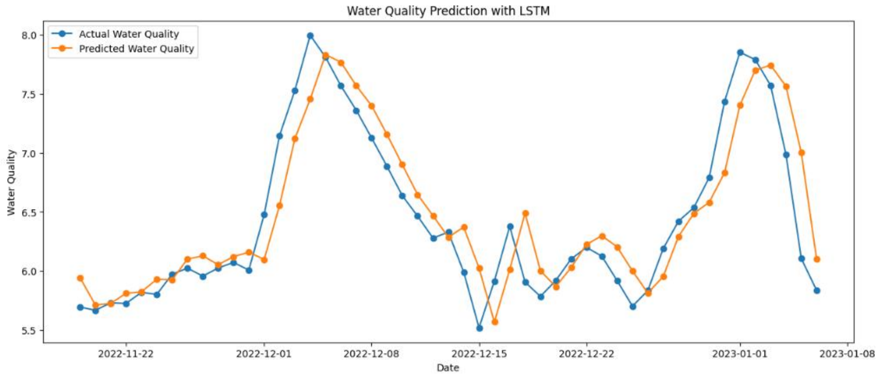


FIGURE 3. *Actual vs Predicted values of LSTM*

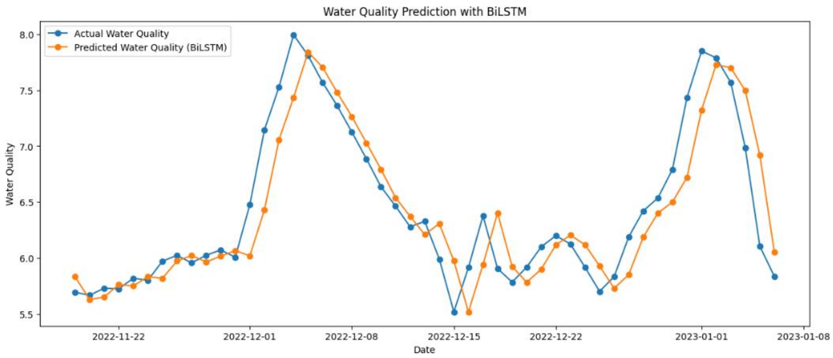


FIGURE 4. *Actual vs Predicted values of BiLSTM*

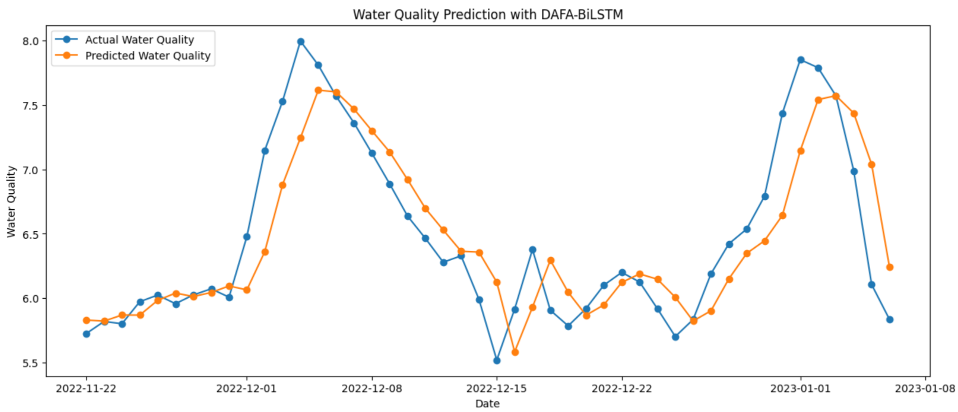


FIGURE 5. *Actual vs Predicted values of DAFA-BiLSTM*

Figure 6 represents a plot showing the training loss and root mean square error (RMSE) over the epochs during the training of the DAFA-BiLSTM model. This plot shows that the RMSE score decreases with the number epochs, implying the model has a better performance.

Journal of Information Systems Engineering and Management
2025, 10(44s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

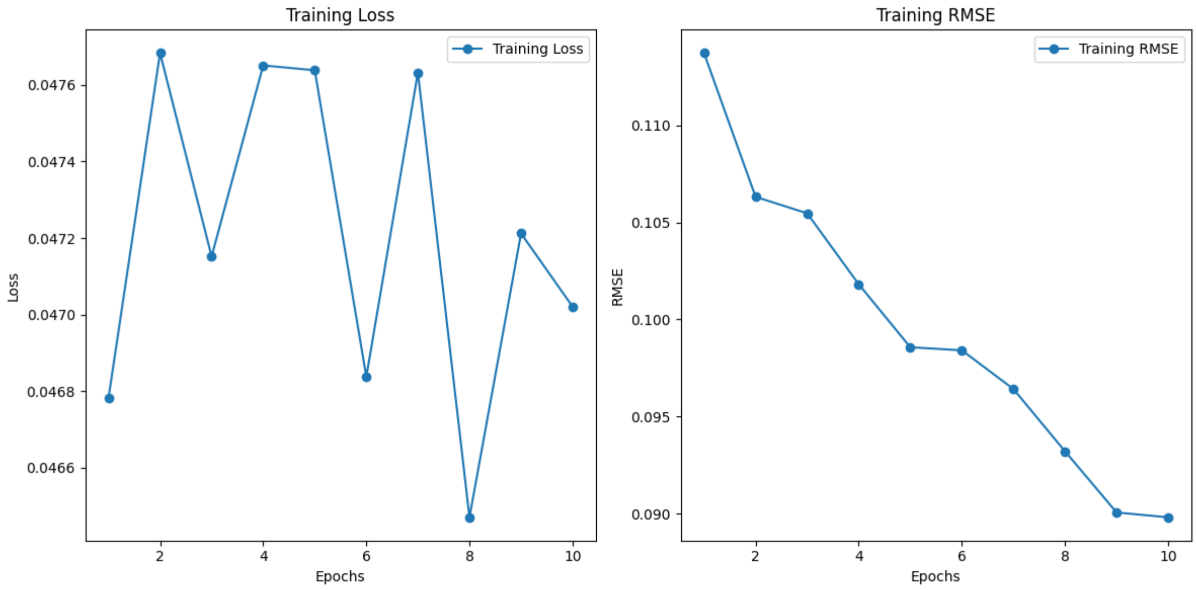


FIGURE 6: Training loss and RMSE vs epochs

On the same test set, the evaluation indicators generated by the AT-LSTM model [1] had the following values: RMSE = 0.405 and R2 = 0.541. On the same test set, however, the evaluation indicators generated by the DAFA-BiLSTM model had the following values: RMSE = 0.216 and R2 = 0.81. This elucidates once more how the combination of deep autoregression and BiLSTM could enhance the DAFA-BiLSTM model's predictive power and accuracy for multivariate time series. The model values are shown in Table 5.2.

Table 5.2: Performance Metrics Comparison of AT-LSTM and DAFA-BiLSTM

Model	RMSE	R2
AT-LSTM	0.405	0.54
DAFA-BiLSTM	0.216	0.81

IV. CONCLUSION

The implementation of the DAFA-BiLSTM model for water quality prediction, particularly in forecasting river water pH levels, shows promising results. The model's architecture, which incorporates attention mechanisms and bidirectional long short-term memory (BiLSTM) networks into autoregression provides various benefits for precisely representing the temporal dynamics present in water quality data. The model does a good job of representing the intricate temporal dependencies found in data on water quality. The model's ability to incorporate additional domain-specific features, such as meteorological data, hydrological parameters, and land use characteristics, enhances its predictive capabilities. The model obtains a more thorough understanding of the factors affecting water quality by taking into account a variety of influencing factors beyond pH values alone. The model's ability to dynamically assess the significance of various time steps in the input sequence is made possible by the addition of an attention mechanism, which increases accuracy. The DAFA-BiLSTM model offers a robust framework for leveraging historical data and domain knowledge to make accurate forecasts. Further research and experimentation may explore additional enhancements to the model architecture and feature engineering techniques to further improve predictive performance and expand its applicability to other water quality parameters and datasets.

Journal of Information Systems Engineering and Management
2025, 10(44s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

References

[1] Chen, H., Yang, J., Fu, X., Zheng, Q., Song, X., Fu, Z., ... & Yang, X. (2022). Water quality prediction based on LSTM and attention mechanism: A case study of the Burnett River, Australia. *Sustainability*, *14*(20), 13231.

[2] Valadkhan, D., Moghaddasi, R., & Mohammadinejad, A. (2022). Groundwater quality prediction based on LSTM RNN: An Iranian experience. *International Journal of Environmental Science and Technology*, *19*(11), 11397-11408.

[3] Hu, Z., Zhang, Y., Zhao, Y., Xie, M., Zhong, J., Tu, Z., & Liu, J. (2019). A water quality prediction method based on the deep LSTM network considering correlation in smart mariculture. *Sensors*, *19*(6), 1420.

[4] Zhang, Q., Wang, R., Qi, Y., & Wen, F. (2022). A watershed water quality prediction model based on attention mechanism and Bi-LSTM. *Environmental Science and Pollution Research*, *29*(50), 75664-75680.

[5] Bao, F., Wu, Y., Li, Z., Li, Y., Liu, L., & Chen, G. (2020). Effect improved for high-dimensional and unbalanced data anomaly detection model based on KNN-SMOTE-LSTM. *Complexity*, *2020*(1), 9084704.

[6] Pyo, J., Pachepsky, Y., Kim, S., Abbas, A., Kim, M., Kwon, Y. S., ... & Cho, K. H. (2023). Long short-term memory models of water quality in inland water environments. *Water research X*, 100207.

[7] Patel, J., Amipara, C., Ahanger, T. A., Ladhva, K., Gupta, R. K., Alsaab, H. O., ... & Ratna, R. (2022). A Machine Learning-Based Water Potability Prediction Model by Using Synthetic Minority Oversampling Technique and Explainable AI. *Computational Intelligence and Neuroscience*, *2022*(1), 9283293.

[8] Ni, Q., Cao, X., Tan, C., Peng, W., & Kang, X. (2023). An improved graph convolutional network with feature and temporal attention for multivariate water quality prediction. *Environmental Science and Pollution Research*, *30*(5), 11516-11529.

[9] Tan, R., Wang, Z., Wu, T., & Wu, J. (2023). A data-driven model for water quality prediction in Tai Lake, China, using secondary modal decomposition with multidimensional external features. *Journal of Hydrology: Regional Studies*, *47*, 101435.

[10] Zhou, J., Wang, Y., Xiao, F., Wang, Y. and Sun, L., 2018. Water quality prediction method based on IGRA and LSTM. *Water*, *10*(9), p.1148.

[11] Tan, W., Zhang, J., Wu, J., Lan, H., Liu, X., Xiao, K., ... & Guo, P. (2022). Application of CNN and long short-term memory network in water quality predicting. *Intelligent Automation & Soft Computing*, *34*(3), 1943-1958.

[12] Lu, J., Zhang, Q., Yang, Z., Tu, M., Lu, J., & Peng, H. (2019). Short-term load forecasting method based on CNN-LSTM hybrid neural network model. *Automation of Electric Power Systems*, *43*(8), 131-137.

[13] Hassan, M. M., Hassan, M. M., Akter, L., Rahman, M. M., Zaman, S., Hasib, K. M., ... & Mollick, S. (2021). Efficient prediction of water quality index (WQI) using machine learning algorithms. *Human-Centric Intelligent Systems*, *1*(3), 86-97.

[14] Ye, Q., Yang, X., Chen, C., & Wang, J. (2019, June). River water quality parameters prediction method based on LSTM-RNN model. In *2019 Chinese control and decision conference (CCDC)* (pp. 3024-3028). IEEE.

[15] Zhou, H., Zhang, Y., Yang, L., Liu, Q., Yan, K., & Du, Y. (2019). Short-term photovoltaic power forecasting based on long short term memory neural network and attention mechanism. *Ieee Access*, *7*, 78063-78074.

[16] Wang, H., Zhang, Y., Liang, J., & Liu, L. (2023). DAFA-BiLSTM: Deep autoregression feature augmented bidirectional LSTM network for time series prediction. *Neural Networks*, *157*, 240-256.