Carbon derived from Tannery Solid Waste LSTM-Based Deep Learning Framework for Real-Time Water Pollution Monitoring in Tannery Effluent–Affected Regions

Conference: 13th AICLST, Kolkata, 6–8 March 2026

Paper reference: TC-26/P

Printed pages: 324–326

Source: *Leather 2050: Challenges & Opportunities*, ILTA (2026)

LSTM-Based Deep Learning Framework for Real-Time Water

Pollution Monitoring in Tannery Effluent–Affected Regions

Srirupa Dasgupta1, Brijesh Kumar Sriwastava2, Ipsita Sengupta3,

3Department of Electronics, 2,5Department of Computer Science and Engineering,

1,4Department of Information Technology Government College of Engineering and Leather Technology, Kolkata-700106, India

Ref. No. TC-26/P

Kalyan Mahata4, Falguni Sinhababu5

Abstract

Water pollution resulting from tannery effluents in the leather industry poses a striking environmental challenge due to high organic load and toxic heavy metals. This study proposes a Long Short-Term Memory (LSTM)–based deep learning framework for real-time monitoring and prediction of water quality. Sequential water quality parameters, including pH, dissolved oxygen, biochemical oxygen demand, chemical oxygen demand, turbidity, and chromium concentration, are modeled as time-series inputs to capture temporal dependencies and pollution dynamics. The LSTM network effectively learns longterm variations and nonlinear patterns in water quality data, enabling accurate pollution level prediction and early anomaly detection. The results highlight the suitability of LSTM models for continuous water pollution surveillance and intelligent environmental management in leather industry–affected regions.

Keywords: tannery effluent, sequential water quality parameters, LSTM, water quality prediction, environmental monitoring

1. Introduction

The leather tanning industry is one of the most environmentally hazardous industrial sectors due to its intensive use of chemicals such as chromium salts, sulfides, acids, and organic compounds. In most cases effluents discharged from tanneries contain high concentrations of

Corresponding author email : srirupa.dasguptam@gmail.com

biochemical oxygen demand (BOD), chemical oxygen demand (COD), suspended solids, turbidity, and toxic heavy metals, particularly chromium. When released untreated or inadequately treated into natural water bodies, these pollutants severely degrade water quality, disrupt aquatic ecosystems, and pose significant health risks to surrounding communities.

Traditional water quality monitoring relies on manual sampling and laboratory analysis, which are costly, timeconsuming, and incapable of providing continuous surveillance. These limitations hinder timely decisionmaking and early pollution control. With the rapid advancement of artificial intelligence and sensor technologies, data-driven models have emerged as powerful tools for environmental monitoring. Among them, Long Short-Term Memory (LSTM) networks [1] are especially suited for modeling time-series data due to their ability to capture long-term dependencies and nonlinear temporal patterns.

This study proposes an LSTM-based deep learning framework for real-time monitoring and prediction of water quality [2]in tannery effluent–affected regions.

2. Methodology

2.1 Data Collection

Water quality data [3] were collected from monitoring stations located downstream of leather tanning industrial

zones. The dataset consisted of regularly sampled measurements of key physicochemical parameters, including pH, dissolved oxygen (DO), biochemical oxygen demand (BOD), chemical oxygen demand (COD), turbidity, and chromium (Cr) concentration. These parameters were selected due to their strong relevance in assessing tannery effluent pollution and regulatory compliance.

2.2 Data Preprocessing

Raw water quality data often contain missing values, noise, and inconsistencies. To address these issues, missing values were treated using linear interpolation. Noise smoothing was performed using moving average filters. All parameters were normalized using Min–Max scaling to ensure uniform feature ranges and improve neural network training efficiency.

The time-series data were then transformed into sequential input windows suitable for LSTM training, enabling the model to learn temporal dependencies between consecutive observations [4].

2.3 LSTM Model Architecture

The proposed deep learning framework employs a multivariate LSTM architecture consisting of:

Table 1: Threshold values Finally, the values of the output parameters obtained from the LSTM are compared with the threshold values to generate an early warning as shown in the sample result shown in Table 2.

An input layer receiving sequential water quality data

Two stacked LSTM layers to capture both short-

term and long-term temporal patterns

Dropout layers to reduce overfitting

A fully connected dense output layer for parameter

prediction

The model was trained using the Adam optimizer with mean squared error (MSE) as the loss function. Backpropagation through time (BPTT) was used to update network weights.

Table 2: Rule-based warning generation

2.4 Model Training and Validation

The dataset was divided into training (70%), validation (15%), and testing (15%) subsets. Model performance was evaluated using standard regression metrics, including Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and the coefficient of determination (R^2). These metrics provided a comprehensive assessment of prediction accuracy and robustness.

3. Results and Discussion

The LSTM-based model demonstrated strong predictive performance across all monitored water quality parameters. The network successfully captured both seasonal variations and abrupt pollution spikes commonly associated with irregular tannery discharge events.

Initially threshold values which serve as regulatory limits are defined using limits set by the local Pollution Control Board as shown in Table 1.

Parameter Threshold

pH 6.0 – 9.0

Turbidity (NTU) 300

TDS (mg/L) 2100

COD (mg/L) 250

Chromium (mg/L) 2.0

Parameter Predicted Limit Status

pH 6.6 6–9 Normal

Turbidity 450 300 ⚠ Warning

TDS 4700 2100 Critical

COD 2250 250 Critical

Chromium 4.2 2.0 □ Critical

Furthermore, the model effectively identified abnormal pollution events when predicted values exceeded regulatory thresholds. This early anomaly detection capability is critical for real-time environmental surveillance and regulatory enforcement in industrial zones.

4. Environmental and Practical Implications

The proposed LSTM-based monitoring system can be integrated with IoT-enabled water quality sensors to form a real-time decision-support platform. Such systems can assist environmental agencies in enforcing pollution control regulations and enable tannery operators to optimize effluent treatment processes. Ultimately, this approach contributes to sustainable industrial practices and protection of aquatic ecosystems.

5. Conclusion

This study presents a deep learning–based LSTM framework for real-time monitoring and prediction of water pollution caused by tannery effluents. By leveraging multivariate timeseries data, the model accurately predicts water quality parameters and detects pollution anomalies

at an early stage. The results demonstrate that LSTM networks are well-suited for intelligent water quality surveillance in leather industry–affected regions.

References

a) Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735–1780.

b) APHA. (2017). Standard Methods for the Examination of Water and Wastewater. American Public Health Association.

c) Zhang, Y., Liu, F., & Wang, X. (2020). Deep learning approaches for water quality prediction: A review. Environmental Modelling & Software, 124, 104585.

d) Singh, K. P., Basant, A., Malik, A., & Jain, G. (2019). Artificial intelligence techniques for water quality assessment and monitoring. Water Research, 152, 111– 123.