HydraSave : A Unique River Water Quality Forecasting and Pollution Control System

Aneena Thankachan
Computer Science and Engineering
Amal Jyothi College of Engineering
Kottayam, India
aneenathankachan22@gmail.com

Devarsh Damodaran
Computer Science and Engineering
Amal Jyothi College of Engineering
Kottayam, India
mymailsdevarsh@gmail.com

Ajay Ben Abraham
Computer Science and Engineering
Amal Jyothi College of Engineering
Kottayam, India
ajaybenabraham18@gmail.com

Arya Shinod
Computer Science and Engineering
Amal Jyothi College of Engineering
Kottayam, India
aryashinod11@gmail.com

Simy Mary Kurian
Computer Science and Engineering
Amal Jyothi College of Engineering
Kottayam, India
simymarykurian@amaljyothi.ac.in

***Abstract*—Technology has contributed a lot to the betterment of the environment, yet there is still a lot to be done in terms of sustaining nature and its resources. Water pollution is one of the most challenging tasks to be harnessed even with present technology. Thus, continuous monitoring and management of water resources is of prime importance. The river and its tributaries are a vital water source for millions of people. Rivers are not only an important source of water but also considered as habitat for animals and plants. Unfortunately, many rivers and streams are significantly polluted all around the world. Therefore, modelling and predicting water quality is of great importance in controlling water pollution. Water is of different quality and hence it can be used for various purposes. The water quality is measured in terms of WQI (Water Quality Index). With the calculated WQI water can be classified into five different grades. This grading allows us to know for what all purposes water of the predicted quality can be used. Even though there are many existing models for predicting water quality, there is still space for improvements as the existing models focuses only on the learning process and also do not perform missing value imputation. To overcome these problems we are proposing a model which is a combination of CNN and Bi-LSTM (Bidirectional Long Short-Term Memory) models for predicting water quality. The preprocessed data is fed into a WQI Generator whose output WQI value is used for predicting WQI for the next year. This predicted WQI is then fed into the LightGBM (light gradient-boosting machine) machine for grading the water. The dataset considered for HydraSave is from the water quality data collected from the Kerala Pollution Control Board (KPCB). HydraSave can be considered as a supporting hand for the pollution control campaigns conducted by the Government of India. Early prediction allows us to take the necessary steps to prevent water from getting highly polluted. Hence, our proposed model acts as a unique river water quality forecasting and pollution control system.**

***Index Terms*—Bi-LSTM, forecasting, LightGBM, pollution control, WQI**

I. Introduction

In all facets of human, environmental, and societal life, water is a fundamental necessity. The fact that water covers 71% of the world has been thoroughly examined. The rivers are significant natural water supplies. There are many uses for river water, including home and agricultural uses. Water is regarded as the most important natural resource on earth among all other natural resources due to its numerous vital uses. Hazardous wastes are, nevertheless, a result of present urbanisation and industry trends. These wastes, which might take the form of metal, organic pollutants, microbiological bacillus, and other things, are directly dumped into the rivers. Human welfare is directly impacted by water quality. The World Water Assessment Programme states that two million tonnes of human waste are disposed into water courses every day. The global rise in health problems and mortality rates, particularly in developing nations, is attributable to the widespread consumption of contaminated water. According to data from developing countries, over 250 million people contract infections annually, resulting in millions of deaths due to diseases in those regions.According to the National Water Quality Monitoring Program (NWQMP), which is run by the Central Pollution Control Board (CPCB) in collaboration with State Pollution Control Boards (SPCBs) and Pollution Control Committees (PPCs), organic pollution is the primary cause of water pollution at 2500 locations across the nation. According to the extent of organic pollution, the CPCB recognised 150 polluted river sections in 2008, and that number rose to 302 in 2015. The discharge of untreated or only partially treated sewage as well as the discharge of industrial wastewater are the main causes of river pollution.

Hence, it is imperative to maintain regular monitoring of water quality to effectively manage its usage across various sectors.Though various models are available for water quality assessment, these needs further improvement on forecasting accuracy and speed. Considering the additional requiremnets for better water quality forecasting we propose a model that combines CNN with Bi-LSTM to provide WQI grading prediction based on sample data followed by grading the water as per its quality using LightGBM technology. The predicted

WQI value helps to analyze for what all purposes the water can be used and what measures can be taken to increase the water quality.

II. LITERATURE SURVEY

Various related works were analyzed from the literature being employed for water quality assessment. Conventional methods for water quality monitoring are often burdened with high implementation costs and time-consuming processes for obtaining results, necessitating further improvements in forecasting and validation. In this regard, machine learning (ML) and deep learning (DL) models have garnered attention for their ability to efficiently learn patterns and offer promising performance.

Sakshi Khullar et al. [1] propose a methodology for assessing water quality in a river using Bi-LSTM network.They trained their model using data collected from river Yamuna in India and tested its performance using various evaluation metrics. The results showed that their proposed model achieved high accuracy in predicting water quality, outperforming other traditional models.The model was tested using monthly water quality data collected at various locations in the Delhi region over a period of 6 years (2013-2019). The findings revealed that the model's predicted values were in strong concurrence with the actual values, demonstrating its ability to accurately forecast future trends. Abhishek Kumar et al. [2] proposed another model in their study, where the methodology section detailed the selection of sampling locations, collection of groundwater samples, and analysis of water quality parameters such as pH, electrical conductivity (EC), total dissolved solids (TDS), hardness, alkalinity, chloride, fluoride, nitrate, sulfate, and heavy metals. The authors emphasized the significance of using the Water Quality Index (WQI) as a decision-making tool for groundwater management and policy formulation. The study revealed that around 20% of the area was classified as non-suitable for drinking water, while the remaining areas fell into categories of good, moderate, poor, and very poor based on the WQI classification. The authors also highlighted that the performance of the Artificial Neural Network (ANN) model, although it contains abundant information about the modeled system, is highly dependent on the training strategy employed.

Xiangqian Wang et al. [5] proposed a hybrid model for gas concentration prediction, utilizing the Long Short-Term Memory (LSTM) network and LightGBM algorithm. Their model incorporated a variable weight combination approach to leverage the strengths of both models and improve prediction accuracy. Real-world gas concentration data was used to evaluate the performance of their proposed model, which demonstrated higher accuracy compared to traditional models. This study provided a novel method for gas accident prevention through gas concentration prediction, although further improvements in performance efficiency may be needed when dealing with rapidly changing input data.On the other hand, HS Atta et al. [9] developed a WQI for assessing the quality of drinking groundwater in the vicinity of the Ismailia Canal in Egypt. They employed a weighted arithmetic mean method to combine the values of various water quality parameters and evaluate the overall groundwater quality. The proposed WQI model was tested on water samples collected from different locations, revealing that the groundwater in the study area was of poor quality and not suitable for drinking purposes. However, it should be noted that Pearson correlation matrix used in the study may not be sufficient to determine nonlinear relationships between variables, and further analysis may be needed to account for such complexities.

SS Baek et al. [10] proposed a novel deep learning model that combined CNN and LSTM networks for the prediction of water level and water quality in rivers. The authors trained their model using data collected from a river in Korea and evaluated its performance using various metrics. The results demonstrated that their proposed model achieved high accuracy in predicting both water level and water quality, surpassing traditional models. However, it should be noted that this study only investigated three pollutants, which is a limited scope compared to existing models. Further research may be needed to expand the scope of pollutants considered in the model.

Md. Galal Uddin et al. [12] presented a comparative study of different WQI models, including their model structures and applications. The study highlighted the parameterization of the models, methods for calculating subindices, parameter weighting values, index aggregation functions, and sources of uncertainty. The WQI model is a popular tool for assessing surface water quality, as it utilizes aggregation techniques to condense large amounts of water quality data into a single value or index. The study also discussed the limitations and issues associated with each model, providing insights into the challenges and considerations in WQI modeling. Yun Ju et al. [6] proposed a combined model of Convolutional Neural Networks (CNN) and LightGBM algorithm for ultra-short-term wind power forecasting. The researchers initially analyzed the properties of raw data from time series of the wind field and adjacent wind field to create new feature sets. Subsequently, CNN was used to extract the feature data, and finally, LightGBM classification was employed to enhance the reliability and accuracy of the forecasting. The proposed model was compared with conventional Support Vector Machine (SVM) models, and it demonstrated higher accuracy in wind power forecasting.

From the analysis of existing methods designed for water quality forecasting, it can be concluded that there is still a need for performance enhancement both in terms of classification efficiency and time complexity.

III. PROPOSED METHODOLOGY

This section provides a detailed description of the proposed methodology of HydraSave: A unique river water quality forecasting and pollution control system. The proposed architecture is shown in fig 1 .The system architecture is designed in such a way that it has less computation cost and more efficiency. Initially, the input data is preprocessed and z-score normalization is performed. This step is done to reduce

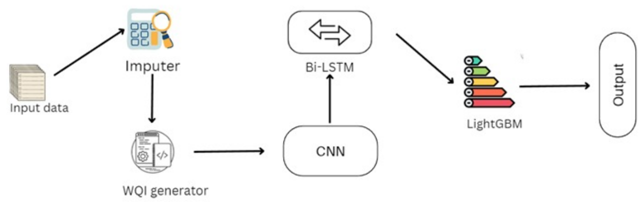

Fig. 1. Proposed architecture of HydraSave

errors and increase accuracy. This data is then fed into the WQI generator which produces WQI values for the complete dataset. The output values are then fed into CNN and Bi-LSTM models, where prediction of new WQI values are done. The predicted WQI values are classified using the LightGBM classifier which inturn classifies the water into any of the five different grades depending on it's quality. With the output grade achieved, the user will be able to understand for what all purposes water of that quality can be used for the predicted year.

A. Input Data

The input data is collected from the Kerala State Pollution Control Board (KSPCB). The data samples of various rivers of Kerala are analyzed. The dataset used in this study includes various water quality parameters such as Dissolved Oxygen (DO), pH, Biochemical Oxygen Demand (BOD), Total Coliform (TC), and Fecal Coliform (FC), which are indicators of water quality. These parameters are used in the calculation WQI. The data thus obtained is converted into a CSV file for further preprocessing. Forecasting the quality of water is crucial as it enables the implementation of necessary pollution control measures and helps in determining the appropriate usage of water based on its quality grade. The input data collected contains water quality parameters of all stations of the river periyar considered. The time series data of the last 14 years from 2007 to 2020 is analyzed. Detailed information about input data is provided in preliminary analysis.

B. Data Preprocessing

Firstly, the data is passed on to an imputer where missing value imputation is done after which normalization is performed. The steps are described below.

1) Imputer : It is important to preprocess the data to avoid errors. This task is performed by the imputer. The missing value imputation technique used in this study involves replacing identified missing values with the mean of neighboring values, using a simple "mean" imputer.
2) Z-Score Normalization : A lot of machine learning algorithms attempt to find patterns in data by comparing features of data points. However, there is an issue when the features are on a completely different scales. Therefore, we are performing z-score normalization.

C. WQI Generator

The water quality index (WQI) is a significant tool used to assess the quality of drinking water in urban, rural, and industrial areas. It is calculated based on the composite influence of various water quality parameters that are considered and weighted in the calculation. The "weighted arithmetic index method" is used to compute WQI, where water quality rating (Qj) is an integral part of the index and is determined using the following expression.

$$Qi = \Sigma qiwi/\Sigma wi - - - - - - - - - - - - (1)$$

Here wi - Unit weight of ith parameter
qi - Quality estimate scale of each parameter, it is calculated with the formula -

$$qi = 100(Vi – VIdeal/Si – VIdeal) - - - - (2)$$

Vi - Measured value of ith parameter
Videal - Ideal value of ith parameter in pure water
Si - Standard value recommended for ith parameter
wi is calculated by the formula -

$$wi = K/Si - - - - - - - - - - - - - - (3)$$

Here K is proportionality constant which is -

$$K = 1/\Sigma Si - - - - - - - - - - - - - - (4)$$

Standard value recommended for parameters (Si)-
DO, mg/L - 10
pH - 8.5
BODd, mg/L - 5
FC/100 mL - 100
TC/100 mL - 1000
Values of unit weights for each parameter (wi) -
DO pH - 0.2604
BOD - 0.4426
FC - 0.0221
TC - 0.0022
Ideal value of parameters (Videal)
DO pH - 7.0
BOD - 0
FC TC - 0
By using the above method, HydraSave calculates the WQI value for each input data.

D. Convolutional Neural Network

Convolution and pooling operations alternately make up the technique known as CNN. Each neuron executes a dot product on a few inputs, followed by an activation function. We will also have a loss or cost function on the final, completely connected layer, much like in the multilayer perceptron. Convolutional neural networks outperform other neural networks when given inputs such as images, voice, or audio. A standard convolutional neural network (CNN) architecture used for prediction tasks typically comprises multiple layers, such as pooling layers, convolutional layers, and fully connected layers. as shown in Fig 2.

Convolutional layers: These layers apply convolutional filters to the input data to extract features. Each filter is a small matrix of weights that slides over the input data, computing a dot product at each position. The result is a feature map that highlights areas of the input that are important for the prediction task. Convolutional layers typically use non-linear activation functions, such as ReLU, to introduce non-linearity into the model.

Pooling layers: These layers downsample the feature maps produced by the convolutional layers. The most common pooling operation is max pooling, which takes the maximum value within a small window and discards the rest. Pooling helps feature map size reduction and make the model more computationally efficient.

Fully connected layers: These layers take the flattened feature maps from the convolutional and pooling layers and use them to make predictions. Fully connected layers are similar to the layers in a traditional neural network, with each neuron receiving input from neurons of previous layer.

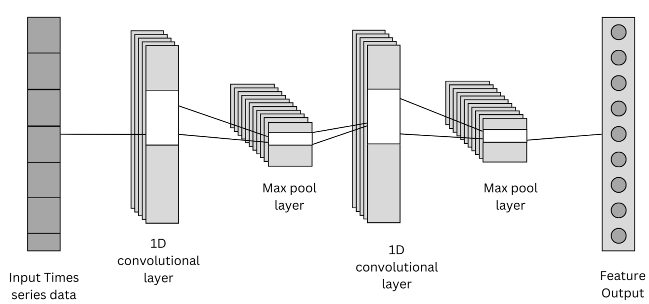

Fig. 2. Structure of CNN

The overall architecture of a CNN for prediction tasks can vary depending on the specific task and dataset.Typically, after the convolutional and pooling layers, one or more fully connected layers are included at the end of the CNN architecture to make predictions. This architecture is often referred to as a ”deep” CNN because of its many layers. Overall, CNN architecture for prediction tasks is highly effective because it allows the model to automatically learn and extract features from input data.

E. Bi-LSTM

Bi-LSTM is a variation of the LSTM network, RNN. Like a standard LSTM, it can process input sequences and maintain an internal state, but it can also process sequences in both directions, making it bidirectional. The input sequence is fed into the BiLSTM layer, which processes it in two directions. That is, forward and backward. The forward direction processes the input sequence from start to end, while the backward direction processes it from end to start. This allows the BiLSTM to capture context from both the past and the future. At each time step, the BiLSTM layer maintains two internal states which includes a forward state and a backward state. The forward flow is updated based on the previous forward state and the current input, while the backward flow is updated based on the previous backward state and the current input. The output of the BiLSTM layer at each time step is the concatenation of the forward and backward states. The output sequence is then passed to the next layer. By incorporating bidirectional processing, the model is designed to analyze inputs in two directions - one from past to future and the other from future to past. This unique approach allows the LSTM to retain information from both the past and future, as the hidden states are combined. This enables the model to effectively capture contextual information from both directions at any given point in time.

F. CNN with Bi-LSTM

In this study, a combined approach of CNN and Bi-LSTM is employed. While CNN is not well-suited for capturing long-term dependencies in time series data due to its lack of internal state, the Bi-LSTM provides a mechanism to address this limitation. CNN is utilized for extracting spatial features from the input time series data, while the Bi-LSTM is used to capture temporal dependencies. The output of the CNN serves as input to the Bi-LSTM, which generates the forecasted output. The CNN model comprises of two one-dimensional convolutional layers followed by two max-pooling layers, which reduces computational complexity. The 1D convolutional layer generates a feature map using ReLU activation function, followed by max-pooling to downsample and select important features. The resulting output is then combined with the Bi-directional LSTM, which processes the data in both forward and backward directions, facilitating the capture of changes in input parameters.

G. LightGBM

LightGBM is a rapid, distributed gradient boosting framework that employs a tree-based learning algorithm. It is renowned for its support of GPU learning, making it a popular choice for data science application development. One of the distinctive features of LightGBM is its leaf-wise splitting approach as shown in 3, where trees are split based on the best fit at the leaf level, in contrast to other boosting algorithms that split trees depth-wise or level-wise. This vertical tree growth strategy sets LightGBM apart, enabling faster training speed, higher efficiency, lower memory usage, improved accuracy, and support for parallel, distributed, and GPU learning. Additionally, LightGBM is capable of handling large-scale data, making it a preferred choice for various applications. XGBoost utilizes a pre-sorted algorithm and a histogram-based approach for optimal split determination, whereas LightGBM employs a groundbreaking technique known as Gradient-based One-Side Sampling (GOSS) to eliminate data instances (examples or observations). In the histogram-based technique, the split value

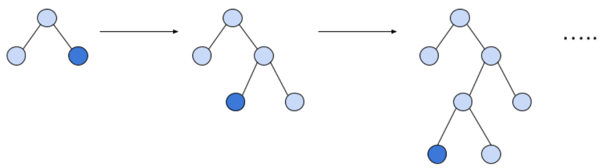

Fig. 3. Level-wise growth and leaf-wise growth

is determined using discrete bins that group all data points for a particular feature. This approach is faster than the pre-sorted method, which considers all potential split points on the pre-sorted feature values. However, LightGBM's GOSS technique is overall faster than the histogram-based approach.

TABLE I
PREPROCESSED DATA AFTER IMPUTATION

WQI Range	Class	Water Quality
0-25	0	Excellent
25-50	1	Good
50-75	2	Medium
75-100	3	Poor
Above 100	4	Unsuitable for drinking

LightGBM classifies the predicted water quality to any of the above given grades in Table I which indicates for what all purposes water of that quality can be used.

IV. PRELIMINARY ANALYSIS

A. Dataset Collection

The concentration of fecal coliform in Periyar, as recorded by the sewage monitoring station, exceeded the allowable limit of 500mpn/100mL, with an average concentration of 113,000mpn/100mL.Thus, it is important to predict the quality of water which helps to know how much polluted the water is.

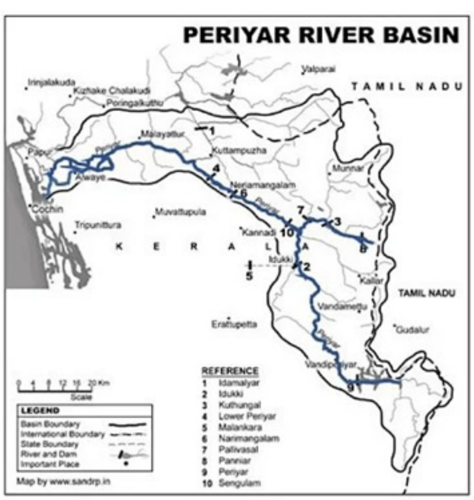

Fig. 4. Map view of Periyar River

The sample data collected contains 5 water quality parameters namely DO, pH, BOD, FC and TC of 8 stations for the last 6 years from 2015 to 2020. The experimental evaluation is conducted by using KSPCB dataset. Several water quality parameters collected from https://kspcb.kerala.gov.in/water-and-air- quality-directory.
The sampling stations are

1) Eloor
2) Kalady
3) KWA Aluva
4) SDP Aluva
5) Pathalam
6) Kalamassery
7) Purappillikkadavu
8) Muppathadam

B. Data preprocessing

1) Imputer: In the dataset, SDP Aluva with station code 1338 is having missing values which is replaced by the mean of it's neighbourhood values.The values before and after imputation is shown in Table II.

TABLE II
PREPROCESSED DATA AFTER IMPUTATION

Parameter	Before	After
Station	1338	1338
DO	NA	6.4
B0D	NA	1.452
pH	NA	6.622
FC	NA	24615.1
TC	NA	28360.83

2) Normalization : For normalization, we used z-score normalization to detect the outliers which are later dropped from the dataset without affecting the performance of the system.

C. Graphical View of The Preprocessed Output Sample

Graphical view of preprocessed output sample is generated to understand and visualize the data after the preprocessed step. The fig 5 Shows the graphical representation of the data.

D. WQI Generation

WQI is computed using the "weighted arithmetic index method". A sample of the generated output is shown in Table III.

V. EXPERIMENT AND RESULT

Initially, the generated WQI values are then fed into CNN. The predicted value is again passed to Bi-LSTM for increasing prediction accuracy.The model predicts the WQI value of the year 2020 of all the 8 stations renamed as S1 to S8 respectively.The below provided graph Fig 6 is plotted to visualize the actual and predicted WQI value of each station respectively.

With the predicted value of WQI, the LightGBM classifies the input sample into any of the five different grades. With

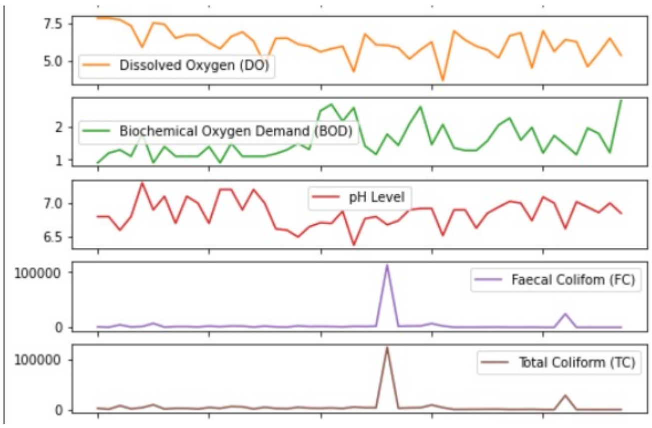

Fig. 5. Graphical View of The Preprocessed Output Sample

TABLE III
SAMPLE DATASET AFTER WQI GENERATION

Station	Location	WQI
18	KALADY	61.187493
3468	KWA ALUVA	47.153553
1338	SDP ALUVA	142.6759
17	ELOOR	99.822425
2333	MUPPATHADAM	57.823608
2334	PATHALAM	202.027134
2336	PURAPPILLIKADAVU	59.639521
2335	KALAMASSERY	75.552203

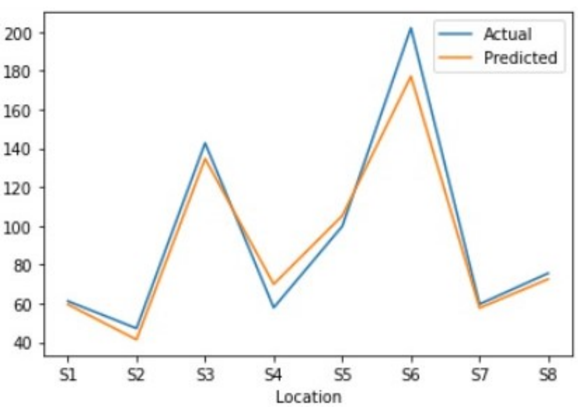

Fig. 6. The graph shows how accurate the result will be for the tested case approximately.

the output grade achieved, the user will be able to understand for what all purposes water of that quality is used.The performance evaluation is done by various measures like precision, recall, F1-Score. Corresponding values are provided in Table IV.

The experimentation result showed that HydraSave requires less computation cost and is much more faster than other existing models. CNN-BiLSTM used in the system showed higher accuracy than CNN-LSTM model. LightGBM classifier has a higher classifying speed when compared to other classification models.Overall, HydraSave is capable of predicting water quality with lesser error rate and higher computation speed.

TABLE IV
PERFORMANCE MEASURE OF LIGHTGBM

Class	Precision	Recall	F1-Score
0	1.00	1.00	1.00
1	1.00	1.00	1.00
2	1.00	0.91	0.95
3	0.91	1.00	0.95
4	1.00	1.00	1.00

VI. FUTURE WORK

Though the system shows much better performance when compared, it still can upgrade its working. Some upgradation and changes are expected to be applied in the system for making it highly efficient. It includes:

1) Rather than considering river Periyar only, the input data will be extended by taking sample data of other rivers of Kerala.
2) Applying transfer learning and thus reduce the overall computation cost.
3) Trying to implement using light-weight models, so that the model can be deployed on a mobile device.

VII. CONCLUSION

HydraSave is a water quality forecasting model that utilizes CNN combined with Bi-LSTM networks. The model explores the applications of Bi-LSTM for forecasting WQI by incorporating data pre-processing, normalization, prediction, and classification of water quality. Traditional CNN and LSTM models often suffer from high computational complexity and low prediction accuracy. To address these limitations, a combination model is used in HydraSave, which helps reduce training errors and improves precision.

Additionally, the HydraSave model incorporates LightGBM technology for water classification based on the predicted WQI value. The model is implemented using time-series data from various monitoring stations of river Periyar in Kerala to study water quality. Accurate prediction of water quality can assist the government in implementing necessary pollution control measures, making HydraSave a unique system for water quality forecasting and pollution control.

REFERENCES

[1] Khullar, S et al. (2022) Water quality assessment of a river using deep learning Bi-LSTM methodology: forecasting and validation. Environ Sci Pollut Res 29, 12875–12889 (2022). https://doi.org/10.1007/s11356-021-13875-w

[2] S.K. et al. (2018) Groundwater Quality assessment using Water Quality Index (WQI) in parts of Varanasi District, Uttar Pradesh, India. J Geol Soc India 92, 76–82. https://doi.org/10.1007/s12594-018-0955-1

[3] S. Chatterjee et.al. (2017), "Water quality prediction: Multi objective genetic algorithm coupled artificial neural network based approach," 2017 IEEE 15th International Conference on Industrial Informatics (INDIN), Emden, Germany, 2017, pp. 963-968, doi: 10.1109/INDIN.2017.8104902.

[4] S. Siami et. al.(2019), "The Performance of LSTM and BiLSTM in Forecasting Time Series," 2019 IEEE International Conference on Big Data (Big Data), Los Angeles, CA, USA, pp. 3285-3292, doi: 10.1109/BigData47090.2019.9005997.

[5] Xiangqian Wang et al. (2022). Prediction of Gas Concentration Based on LSTM-LightGBM Variable Weight Combination Model Energies 15, no. 3: 827. https://doi.org/10.3390/en15030827

[6] Y. Ju et al. (2019). A Model Combining Convolutional Neural Network and LightGBM Algorithm for Ultra-Short-Term Wind Power Forecasting, in IEEE Access, vol. 7, pp. 28309-28318, 2019, doi: 10.1109/ACCESS.2019.2901920.

[7] Şehnaz et al. (2017). Evaluation of water quality using water quality index (WQI) method and GIS in Aksu River (SW-Turkey), Science of The Total Environment,Volumes 584–585,Pages 131-144,ISSN 0048-9697, https://doi.org/10.1016/j.scitotenv.2017.01.102.

[8] Md. Saikat Islam Khan et al. (2021), Water quality prediction and classification based on principal component regression and gradient boosting classifier approach, Journal of King Saud University - Computer and Information Sciences, Volume 34, Issue 8, Part A, September 2022, Pages 4773-4781
https://www.sciencedirect.com/science/article/pii/S1319157821001361

[9] HS Atta et al. (2022), Water quality index for assessment of drinking groundwater purpose case study: area surrounding Ismailia Canal, Egypt, Journal of Engineering and Applied Science , 69,1315
https://jeas.springeropen.com/articles/10.1186/s44147-022-00138-9

[10] SS Baek et al. (2020), Prediction of Water Level and Water Quality Using a CNN-LSTM Combined Deep Learning Approach, Water 2020, 12,3399
https://www.mdpi.com/2073-4441/12/12/3399

[11] A. K. Kadam et al. (2019) Prediction of water quality index using artificial neural network and multiple linear regression modelling approach in Shivganga River basin, India. Model. Earth Syst. Environ. 5, 951–962 https://doi.org/10.1007/s40808-019-00581-3

[12] Md. Galal Uddin et al (2021) A review of water quality index models and their use for assessing surface water quality, Ecological Indicators, Volume 122, 107218, ISSN 1470-160X, https://doi.org/10.1016/j.ecolind.2020.107218

[13] Dao Nguyen Kho et al. (2022) Using Machine Learning Models for Predicting the Water Quality Index in the La Buong River, Vietnam. Water 14,no. 10:1552. https://doi.org/10.3390/w14101552

[14] Shahid, Farah et al (2020). Predictions for COVID-19 with deep learning models of LSTM, GRU, and Bi-LSTM. Chaos, Solitons & Fractals, Volume 140, 2020, 110212, ISSN 0960-0779, https://doi.org/10.1016/j.chaos.2020.110212.