Base Paper Review

Name : Water Quality Prediction Based on Machine Learning Techniques

Link : Paper Link

The paper is based on the Adaptive Neuro Fuzzy Inference System (ANFIS) many machine learning techniques, like multivariate linear regression (MLR) and artificial neural network (ANN) model, have been proposed to address the problem. However, simple linear regression analysis cannot accurately forecast water quality because of complicated linear and nonlinear relationships in the water quality dataset.

The ANN model also has shortcomings though it can accurately predict water quality in some scenarios. Artificial neural networks (ANN) are adopted to explore the non-linear relationships residing in water quality datasets. Various ANN models have been designed to predict water and wastewater discharge quality based on previous existing datasets. A comprehensive comparison between ANN and MLR models for oxygen demand prediction has been performed. The experimental results show that a three-layer neural network model outperforms an MLR model. In, neural network models are used to predict four parameters and the proposed model has higher accuracy and better stability in the experiment. The Adaptive Neuro-Fuzzy Inference System (ANFIS) has been proven to be an effective tool in formulating the complicated linear and non-linear relationship hidden in datasets.

Although the ANFIS model can achieve good performance in the water quality prediction, it has some limitations. Firstly, the size of training dataset should not be less than the number of training parameters required in the model. Secondly, when the data distribution in the testing dataset is not reflected in the training dataset, the ANFIS model may generate out-of-range errors.

Name : River Water Quality Prediction and index classification using Machine Learning

Link : Paper Link

The idea of the paper is to explore the methodologies that have been employed to help solve problems related to water quality. Typically, conventional lab analysis and statistical analysis are used in research to aid in determining water quality, while some analyses employ machine learning methodologies to assist in finding an optimized solution for the water quality problem.

In this paper, the proposed methodology employs four input parameters, namely, temperature, turbidity, pH, and total dissolved solids. Of all the employed algorithms, gradient boosting, with a learning rate of 0.1, and polynomial regression, with a degree of 2, predict the WQI most efficiently, having a mean absolute error (MAE) less than 3, respectively. Whereas multi-layer perceptron (MLP), with a configuration, classifies the WQC most efficiently, with an accuracy of nearly 85%. The proposed methodology achieves reasonable accuracy using a minimal number of parameters to validate the possibility of its use in real-time water quality detection systems.

This paper explored an alternative method of machine learning to predict water quality using minimal and easily available water quality parameters. A set of representative supervised machine learning algorithms were employed to estimate WQI. This showed that polynomial regression with a degree of 2, and gradient boosting, with a learning rate of 0.1, outperformed other regression algorithms by predicting WQI most efficiently, while MLP outperformed other classification algorithms by classifying WQC most efficiently.

Name : Efficient Water Quality Prediction Using Supervised Machine Learning

Link : Paper Link

The water quality index is determined by 27different parameters affecting water quality like dissolved oxygen, temperature, pH, alkalinity, hardness, chloride, coliform, etc. Data normalization and feature selection are done to construct the dataset to develop machine learning models. Machine learning algorithms such as linear regression, MLP regressor, support vector regressor, and random forest have been employed to build a water quality prediction model. Support vector machines (SVM), naïve Bayes, decision trees, and MLP classifiers, have been used to develop a classification model for classifying water quality index. The experimental results revealed that the MLP regressor efficiently predicts the water Quality index with root mean squared error of nearly 2.5, MLP classifier classifies the water quality index with 81% accuracy. The developed models show promising output concerning water quality index prediction and classification.

The water quality prediction model is developed by learning the trends in the river water quality dataset using linear regression, random forest, support vector regression, and multilayer perceptron regression. Here water quality index is the target variable and all other parameters in the dataset are independent variables in modelling the regression process. Regression can be used to forecast a response using a new set of predictors.

It is concluded that, MLP regressor and MLP classifier out performs than the other models in forecasting water quality index. In future, hybrid models with deep learning algorithm can be built to improve the efficiency of the water quality prediction.