Chapter 4
Prediction of Water Quality Using Machine Learning

Tran Thi Hong Ngoc
An Giang University. Vietnam

Phan Truong Khanh
An Giang University, Vietnam

Sabyasachi Pramanik
https://orcid.org/0000-0002-9431-8751
Haldia Institute of Technology, India

ABSTRACT

With the fast growth of aquatic data, machine learning is essential for data analysis, categorization, and prediction. Data-driven models using machine learning may effectively handle complicated nonlinear problems in water research, unlike conventional approaches. Machine learning models and findings have been used to build, monitor, simulate, evaluate, and optimize water treatment and management systems in water environment research. Machine learning may also enhance water quality, pollution control, and watershed ecosystem security. This chapter discusses how ML approaches were used to assess water quality in surface, ground, drinking, sewage, and ocean. The authors also suggest potential machine learning applications in aquatic situations.

DOI: 10.4018/979-8-3693-1396-1.ch004

1. INTRODUCTION

Wastewater carrying toxins from fast economic growth threatens natural water ecology. So, several water pollution management methods evolved. Water quality analysis and assessment have greatly enhanced water pollution control efficiency. The multivariate statistical approach, fuzzy inference, and water quality index (WQI) are among the various methods used to monitor and measure water quality globally. While many water quality metrics may be traced in accordance with regulations, the final findings may vary according to parameter selection. Taking into account that all water quality metrics is impractical due to cost, technical difficulty, and inability to account for variability. Recently, developments in machine learning have led academics to anticipate that huge volumes of data may be achieved and assessed to accomplish complex and large-scale water quality monitoring needs.

ML algorithms are used in artificial intelligence to examine data and find patterns to forecast future information. With its accuracy, flexibility, and extensibility, machine learning has become a popular data analysis and processing tool in several fields. Machine learning simplifies the finding of underlying mechanisms for complex nonlinear relational data. Recently, ML showed a huge potential as a tool in ecological science and engineering due to its versatility. In spite of the difficulty of ML for water quality measurement and assessment, much precise results are predicted.

Complex water kinds include drinking, wastewater, and groundwater, surface, marine, and fresh. These water kinds have varied qualities, making quality study difficult. Previous research suggests that machine learning may effectively handle these difficulties. In this study, we address the pros and cons of typical ML approaches and their implementations and performance in surface water, groundwater, drinking water, wastewater, and ocean water (Fig. 1).

ML is commonly utilized to find insights or have predictions from vast data from many contexts. Prior to using ML, data collecting, algorithm selection, model training, and validation are needed. Among these methods, selecting algorithm is the key aspect.

Machine learning has two primary classes: supervised and unsupervised. Labels in datasets distinguish these two kinds. Supervised learning predicts from labeled training datasets. Input and anticipated output values are included in each training instance. Supervised learning algorithms discover input-output correlations and create a predictive model to estimate the outcome from the I/P data. Supervised learning methods, such as LR, ANN, decision trees, SVM, Naive Bayes, KNN, and random forests are designed for data classification and regression.

In contrast, unsupervised learning handles data generally without labels, addressing pattern recognition problems using unlabeled training datasets. Unsupervised learning classifies training data depending on features, primarily via dimensionality reduction

Related Content

Impact of Climate Change on Groundwater Resources
C. P. Kumar (2017). *Natural Resources Management: Concepts, Methodologies, Tools, and Applications (pp. 1094-1120).*
www.irma-international.org/chapter/impact-of-climate-change-on-groundwater-resources/165337

Atmospheric Processes and Climate Change
Jaya Yadavand Dyvavani Krishna Kapuganti (2024). *Advanced Geospatial Practices in Natural Environment Resource Management (pp. 1-25).*
www.irma-international.org/chapter/atmospheric-processes-and-climate-change/342208

Farming Adaptations to the Impacts of Climate Change and Extreme Events in Pacific Island Countries: Case Study of Bellona Atoll, Solomon Islands
Viliamu Iese, Joseph Maeke, Elisabeth Holland, Morgan Wairiuand Sumeet Naidu (2017). *Natural Resources Management: Concepts, Methodologies, Tools, and Applications (pp. 852-875).*
www.irma-international.org/chapter/farming-adaptations-to-the-impacts-of-climate-change-and-extreme-events-in-pacific-island-countries/165324

Land Surface Temperature and Its Lapse Rate Estimation Using Landsat-8 TIRS Data in Beas River Basin, India and Computed Differences With MODIS-Terra
Gopinadh Rongali, Ashok K. Keshari, Ashwani K. Gosain, R. Khosaand Ashish Kumar (2024). *Advanced Geospatial Practices in Natural Environment Resource Management (pp. 129-148).*
www.irma-international.org/chapter/land-surface-temperature-and-its-lapse-rate-estimation-using-landsat-8-tirs-data-in-beas-river-basin-india-and-computed-differences-with-modis-terra/342214

Optimization of Natural Gas Liquefaction Process
Mohd Shariq Khanand Moonyong Lee (2017). *Natural Resources Management: Concepts, Methodologies, Tools, and Applications (pp. 432-456).*
www.irma-international.org/chapter/optimization-of-natural-gas-liquefaction-process/165305