

Performance analysis of different machine learning models on groundwater quality prediction for sustainable groundwater management

Annie Jose* and Srinivas Yasala

Centre for Geotechnology, Manonmaniam Sundaranar University, Tirunelveli 627 012, India

Both natural and man-made environmental factors affect the chemical composition of groundwater. A large amount of chemical data is analysed to determine if groundwater is suitable for drinking and other uses. Many machine learning methods are available for evaluating groundwater quality. Using machine learning is considered convenient and cost-effective for analysing large datasets and saving time. The present study compares the performance of logistic regression, k-nearest neighbours, and support vector machine to identify the most suitable algorithm for assessing groundwater quality. The present study area is the Kanyakumari district, Tamil Nadu, India.

Keywords: Groundwater quality, Kanyakumari district, machine learning algorithms, prediction, safe drinking water.

using physical parameters effectively represent GWQ for irrigation at a lower cost in Northwest Morocco. Sahoo *et al.*⁶ focused on utilising ML algorithms to model changes in groundwater levels in agricultural regions, particularly in the High Plains' aquifer and the Mississippi River Valley alluvial aquifer in the U.S. The present study introduced an ensemble modelling framework combining spectral analysis, ML techniques, and uncertainty analysis to predict groundwater level changes. To estimate the water quality index in Haryana, Raheja *et al.*⁷ used ML models, including deep neural networks, XG boost, and gradient boosting, to identify key parameters influencing GWQ. Numerous studies employing ML approaches have been conducted globally to predict water quality. A thorough understanding of the past, present, and future of GWQ can help policy makers and practitioners in the water sector develop strategies for planning and managing water resources, ensuring sustainable water development. Machine learning methods, alongside traditional approaches, help uncover complex data patterns and improve GWQ predictions⁸.