ABSTRACT

This study builds a new predictor based on the intelligent computational representation by supervised learning techniques (i.e., MARS) and optimization algorithm (i.e., Bat) that miner calls ($IM^{12}CP$-WQI). The main goal of that miner is to find water quality based on 12 types of concentrations that cause water pollution, which are: potential hydrogen (PH), total dissolved solids (TDS), turbidity unit (NTU), total hardness (TH), total alkalinity, calcium (Ca), magnesium (Mg), potassium (K), sodium (Na), chloride (Cl), nitrate (NO_3), and sulfate (SO_4). $IM^{12}CP$-WQI is multi-phased; the first phase relates to collection of data through two seasons (i.e., summer and winter). The second phase, called preprocessing of data, includes (a) combination samples related to each season with its focus independently and (b) normalization the dataset through MinMaxScaler to make dataset in range (0, 1). The third phase is related to building the predictor based on developing MARS through replaced the kernel function of it (i.e., base function) by one of optimization algorithm called (i.e., Bat) to find the best of weight for each concentration, number of base model "M" of that MARS. After that, dividing the dataset based on five cross-validation principles into training and testing dataset, we will use the training dataset to build the predictor. While testing dataset evolution in the final phase, the results will offer an evaluation of each season by taking the readings of all concentrations for reigns and then computing three measures called R2, NSE, and D. We can summarize the main stages of this research as explained below: collection the concentrations of the PH, TDS, NTU, TH, TA, Ca, Mg, K, Na, Cl, NO_3, and SO_4 by dealing with a dataset of each season. Splitting the dataset of each season into parts, training, and testing the dataset based on five cross-validation concepts, we build a new predictor called Intelligent Miner based on twelve concentrations to predict water quality index ($IM^{12}CP$-WQI) by the combination of the most advantageous features of both MARS and BA. Evaluation of the results of $IM^{12}CP$-WQI is based on three measures.