Machine learning-based system for supporting ecological status assessment of lake ecosystems

Abstract

The idea of protecting lake ecosystems has solidified itself as one of the most crucial environmental issues for the European society since the Water Framework Directive was ratified in 2000. Additionally, the evaluation of their quality allowed for a formal review to see if the Member States' concurrent socioeconomic development is hampering the inland waters' desired ecological condition. The major objective of the study was to create a machine learning algorithm-based support system for systematic assessment. In order to meet the demands of supporting decision-making processes during evaluation, prioritization of protective actions, and monitoring environmental target achievement, ecological indicators of state must be automated and repeatable. The system is made up of three interconnected parts that propose solutions in the areas of ecosystem prioritizing, process modeling, and missing data imputation. For calibration and testing, data from Poland's 499 lakes from the second planning cycle of revising the Water Management Plans (2010-2015) were used. A solution to the issues relating to the shortcomings in the data utilized to create the ecological assessment indicators was put forth in the system's initial phase. Data gaps force the abandonment of evaluation or the employment of imputation techniques that are chosen only for their availability and acceptance. The suggested approach is based on a fusion of hierarchical clustering and machine learning methods. This gave the chance to fill the gaps left by the semi-supervised imputation procedure. Emphasis was also made on the interaction of the ecological assessment expert with the data analyst in order to collect the optimal set of data to develop the assessment model at an early stage of work with data. An ecological state classifier based on a support vector machine and a radial basis function kernel (kSVM) lies at the heart of the system. Based on the findings of the completed assessment and measurement information for the factors that were used to build indicators of the ecological state of lakes, the categorization was made. Using a mix of linear discriminants, the classification was performed in a space with two leading dimensions. This allowed for a replication of the expert analysis that was more than 94% accurate. By optimizing the kSVM classifier's hyperparameters, the constructed model was made resilient to changes in the distribution of the input variables. A tool supporting decision on the prioritizing of the examined lakes within ecological status classes was developed using the model that was derived. The method created scenarios of prioritizing outcomes, which were then translated to real classification results given by experts, and utilized the unsupervised Kohonen algorithm to group lakes. This allowed locations where subgroups emerge within one of the classes of ecological state, which could not be detected in the initial set of evaluations, to be found. Information was given to the lakes in each of the newly formed clusters regarding their placement in reference to the original categorization and in potential division scenarios. This process allowed for the priority to occur while retaining the ecological status class structure mandated by the WFD's rules. At each level of working with data, recommendations for data visualization have been built into the system. In addition to between the expert and the audience, visual communication is essential. The proper technique of presenting the outcomes of later phases may substantially aid in the understanding of even the most complicated information when it comes to interactions within ecological status assessment teams. Therefore, the major target for this work is specialists in assessing the ecological state of inland surface waters, experts in putting the WFD's requirements into practice, and the data analysts who assist them.

Keywords: decision support systems, machine learning, lake ecological state