

Summarizing Data Sets for Data Mining by Using Statistical Methods in Coastal Engineering

Authors : Yunus Doğan, Ahmet Durap

Abstract : Coastal regions are the one of the most commonly used places by the natural balance and the growing population. In coastal engineering, the most valuable data is wave behaviors. The amount of this data becomes very big because of observations that take place for periods of hours, days and months. In this study, some statistical methods such as the wave spectrum analysis methods and the standard statistical methods have been used. The goal of this study is the discovery profiles of the different coast areas by using these statistical methods, and thus, obtaining an instance based data set from the big data to analysis by using data mining algorithms. In the experimental studies, the six sample data sets about the wave behaviors obtained by 20 minutes of observations from Mersin Bay in Turkey and converted to an instance based form, while different clustering techniques in data mining algorithms were used to discover similar coastal places. Moreover, this study discusses that this summarization approach can be used in other branches collecting big data such as medicine.

Keywords : clustering algorithms, coastal engineering, data mining, data summarization, statistical methods

Conference Title : ICDMBDDDS 2017 : International Conference on Data Mining, Big Data, Database and Data System

Conference Location : Toronto, Canada

Conference Dates : June 15-16, 2017