THE EDITED NEAREST NEIGHBOR RULE BASED ON THE REDUCED REFERENCE SET AND THE CONSISTENCY CRITERION

Marcin Raniszewski

Technical University of Łódź, Computer Engineering Department, Łódź, Poland

Abstract

In this paper a new editing procedure for the Nearest Neighbor Rule (NN) is presented. The representativeness measure is introduced and used to choose the most representative samples of the classes. These samples constitute a reduced reference set. An edited reference set is created from all the training set samples (including samples from the reduced set), which are correctly classified by the NN rule operating with the reduced set. The performance of the presented method is evaluated and compared with five other well-known editing techniques, on five medical datasets.

Keywords: editing techniques, nearest neighbor rule, consistency criterion, reference set reduction, representativeness measure