

DOUBLE SORT ALGORITHM RESULTING IN REFERENCE SET OF THE DESIRED SIZE

Marcin Raniszewski

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

Abstract

An algorithm for obtaining the reduced reference set that does not exceed the desired size is presented. It consists in double sorting of the original reference set samples.

The first sort key of the sample x is the number of such samples from the same class, that sample x is their nearest neighbour, while the second one is mutual distance measure proposed by Gowda and Krishna. The five medical datasets are used to compare the proposed procedure with the RMHC-P algorithm introduced by Skalak and the Gowda and Krishna algorithm, which are known as the most effective ones.

Keywords: reference set reduction, mutual distance measure, representative measure, Gowda and Krishna algorithm, Hart's algorithm, Skalak's RMHC-P algorithm, Double Sort Algorithm