

Abstract

Title: Similarity Measures of Interval-Valued Fuzzy Sets in Classification of Uncertain Data. Applications in Ovarian Tumor Diagnosis

The dissertation deals with the problem of measuring the similarity when knowledge about objects represented by the Interval-Valued Fuzzy Sets is incomplete and uncertain. Various approaches to measuring similarity of classical and interval-valued fuzzy sets were investigated and compared. It appears that to be able to take full account of the data incompleteness, it is necessary to express the similarity as an interval. Theory necessary to properly model interval similarity was built. Basic properties, which in this case should be fulfilled by similarity measure were formulated, and a construction method of infinitely many such measures was proposed. This method allows to construct a new interval measure from a similarity measure of fuzzy sets, as long as it meets certain conditions. Problem of effective calculation of the new measures obtained by this method was examined. Special attention was given to the generalized version of the Jaccard Index. Using the interval similarity measures, two classification methods that allow full support for data uncertainty, both at the stage of building a classifier and its usage, were proposed. Comprehensive evaluation of the classification quality using real medical data was performed. One of the proposed methods was applied in the intelligent diagnosis support system for Ovarian Tumor – OvaExpert.

Keywords: fuzzy sets, similarity, uncertainty, classification.

Patryk Żywica