MSc Sylwia Buregwa-Czuma Rzeszow, 2017-05-18

Interdisciplinary Centre for Computational Modelling

Faculty of Mathematics and Natural Sciences

University of Rzeszow, Poland

Title of the dissertation:

„Methods of domain knowledge application to improve the quality of classifiers"

Abstract

The dissertation deals with methods that allow the use of domain knowledge to improve the quality of classifiers, where quality improvement concerns: feature extraction methods, classifier construction methods, and methods for predicting decision values for new objects. In particular the following methods have been proposed to improve the quality of classifiers: the expert features (attributes) defined using domain knowledge expressed in a language that uses the temporal logic, a new method of measuring the quality of cuts during supervised discretization using a matrix of the distances between decision attribute values defined by a domain knowledge, a new decision tree that uses redundant cuts to verify the partition of a tree node, a new method for determination of similarities between objects (e.g. patients) using an ontology defined by an expert with its application to the k-nearest neighbors classifier construction and a new method for generating cross rules describing the effect of a factor interfering perception based on a classifier.

All of the aforementioned methods have been implemented in the CommoDM software library, which is one of the RSES-lib library extensions.

Implemented methods have been tested on real data sets. These were comparative data sets known from the literature as well as own medical data sets collected during the preparation of the dissertation. The latter data sets are associated with the medical aspect of the dissertation that deals with the support of treatment of patients with stable ischemic heart disease, and the main medical problem considered in the thesis is the problem of predicting the presence of significant coronary artery stenosis based on non-invasive heart monitoring by Holter method.

The results of experiments confirm the effectiveness of the application of additional domain knowledge in the task of creating and testing classifiers, because after the application of new methods the quality of classifiers has increased considerably. At the same time, the clinical interpretation of the results is more consistent with medical knowledge.

The research has been supported by the grant DEC-2013/09/B/ST6/01568 and the grant DEC-2013/09/B/NZ5/00758, both from the National Science Centre of the Republic of Poland. Their results were published in 10 publications, including 3 publications in journals from the A list of the Polish Ministry of Science and Higher Education, 3 publications indexed in the Web of Science, one chapter in a monograph and 3 post-conference publications.