

RSKT'2008 Tutorial Proposal: Rough Set Approach to KDD

ANDRZEJ SKOWRON, HUNG SON NGUYEN

Institute of Mathematics, Warsaw University

Emails: {skowron,son}@mimuw.edu.pl

Abstract

This tutorial aims to revise the fundamentals of Rough Set theory and its application in Knowledge Discovery from Databases (KDD). The practical guide to analysis of different real-life problems using rough set methods as well as the presentation of Rough Set Exploration System (RSES) can be treated as the preliminary material for the main conference and associated workshops.

1 Motivations

Rough set theory was introduced by Zdzisław Pawlak in the early 80's and has currently reached a level of high visibility and maturity. Originally, Rough Sets, whose main philosophy is based simply on discernibility between objects, were presented as an approach to concept approximation under uncertainty. This brilliantly simple idea has been successively expanded in the last twenty years. Many effective methods for data analysis have been developed on the basis of rough set theory.

In recent years, a growth of interest in rough set theory and its applications can be seen in the number of research papers submitted to international workshops, conferences, journals and edited books, including two main biannual conferences on rough sets and the special sub-line of LNCS series. A large number of efficient applications of rough sets in Knowledge Discovery from various types of databases have been developed. Rough sets are applied in many domains, such as, e.g., instance, medicine, finance, marketing, telecommunication, conflict resolution, text mining, intelligent agents, image analysis, pattern recognition, bioinformatics.

This tutorial is intended to fulfill the needs of many researchers to understand the rough set methodologies in mining of standard and nonstandard data. The methodology based on rough sets can serve as a useful complementary tool for other data mining methods. The tutorial should help the audience to find out if some of the presented methods may support their own KDD or/and DM research.

2 Detailed outline

The tutorial is intended to occupy four slots of 45 minutes each.

- Fundamentals of rough set theory;
- Rough set approach to KDD;
- Rough set methods: data reduction, rule extraction, discretization, decomposition, hierarchical learning, ...;
- Practical guide for Rough Set Exploration System (RSES);
- Some exemplary applications of RSES (with exercises);
- Emergent Challenging Problems.

3 Potential target audience

The first part of this tutorial is targeted to those participants who want to learn rough set theory by examples at the basic level in the context of KDD problems. Only elementary knowledge on the typical data analysis tasks is assumed. No knowledge on rough sets and related paradigms is required. The second part of this tutorial is also targeted to young researchers who would like to apply the rough set approach to the real-life problem solving.