Comparative analysis of data mining tools application to industrial data generated from PLM ontology

Spring 2013

Keywords: Data Mining, Industrial Data, Data Mining Tool Applicability, Automatic Data Analysis

Problem: One of the biggest challenges in creating automatic data mining tool is methodology selection and result validation. In data mining processes, depending on a nature, amount, data types present and results expected, human expert has to decide on algorithm to be used for pattern recognition. Ontology gives us advantage of having clean, structured data with clearly defined relations. Having this, as well as the experience that real-life industrial data are not expected to have very complex patterns, we need to predefine potential courses of automatic data mining procedure, depending only on data characteristics.

Plan: 1. Study the most popular data mining algorithms
    2. Study industrial data present in the ontology
    3. Analyze applicability of algorithms to different data types, data set sizes and data quality
    4. Study common domain of application for different algorithms

Supervisor: Dr. Dimitris Kiritsis, (dimitris.kiritsis@epfl.ch)

Responsible collaborator(s): Dr. Soumaya El Kadiri (soumaya.elkadiri@epfl.ch)
    Ana Milicic (ana.milicic@epfl.ch)

Duration: 4 Months