Abstract

In Data Analysis, Data Mining, and Machine Learning, discovery of knowledge is an important issue. The feature Subset Selection is a prominent problem within Machine Learning and Data Mining. Rough Set Theory plays a major role in knowledge discovery and several works have emphasized the role of the basic constructs of rough set approach in feature selection, namely Reducts and their approximations. Rough Set Theory plays a major role in knowledge discovery and it is mainly used in the classification and knowledge reduction tasks. Scalable algorithms for feature selection using Rough Set Theory is gaining prominence and finding set of Reducts based on Discernibility Matrix, is one alternative solution provided the corresponding algorithms improve on three points of view which are a) Computation time, b) the Reducts size and c) the accuracy of the induced model.

The core of the thesis, deals with the process of finding out the most Stable Reducts in the Object-Oriented Rough Set Model by designing the corresponding techniques for Dynamic Reduct, Core of Object-Oriented Rough Set Model and Sub-Object-Oriented Rough Set Model, the(F-λ)-Dynamic Core, Generalized Dynamic Core. The work is then extended to compute stable Reducts in a series of Sub-Object-Oriented Rough Set Models, which are named Parallel Reduct while focusing on the need to have them. The thesis, also deals
with another new Reduct called Relative Reduct along with algorithms for finding them. Finally this thesis addresses some issues relating to intense computation of Discernibility Matrix and positive regions and proposes a concept of Relative Class Dependency to fix some issues. The work is verified by the application of examples to examine the results and to determine whether the techniques proposed are effective and efficient with regard to accuracy for solving the feature reduction problems.