Abstract

In general rule induction algorithms have a bias that favors the discovery of large disjuncts, rather than small disjuncts. In the context of data mining, small disjuncts are rules covering a small number of examples. Due to their nature, small disjuncts are error prone. It correctly classify individually only few examples but, collectively, cover a significant percentage of the set of examples, so that it is important to develop new approaches to cope with the problem of small disjuncts. This paper presents a classification algorithm based on Evolutionary Algorithm (EA) that discovers interesting small-disjunct rules in the form If P Then D. The proposed system specifically designed for discovering rules covering examples belonging to small disjuncts. The proposed algorithm is validated on several datasets of UCI data set repository and the experimental results are presented to demonstrate the effectiveness of the proposed scheme for automated small-disjunct rules mining.

References


Index Terms

Computer Science Pattern Recognition

Keywords

Interestingness Small Disjunct Predictive Accuracy Genetic Algorithm