

Novak, Petra Kralj; Lavrač, Nada; Webb, Geoffrey I.

Supervised descriptive rule discovery: a unifying survey of contrast set, emerging pattern and subgroup mining. (English) Zbl 1235.68178

J. Mach. Learn. Res. 10, 377-403 (2009).

Summary: This paper gives a survey of contrast set mining (CSM), emerging pattern mining (EPM), and subgroup discovery (SD) in a unifying framework named supervised descriptive rule discovery. While all these research areas aim at discovering patterns in the form of rules induced from labeled data, they use different terminology and task definitions, claim to have different goals, claim to use different rule learning heuristics, and use different means for selecting subsets of induced patterns. This paper contributes a novel understanding of these subareas of data mining by presenting a unified terminology, by explaining the apparent differences between the learning tasks as variants of a unique supervised descriptive rule discovery task and by exploring the apparent differences between the approaches. It also shows that various rule learning heuristics used in CSM, EPM and SD algorithms all aim at optimizing a trade off between rule coverage and precision. The commonalities (and differences) between the approaches are showcased on a selection of best known variants of CSM, EPM and SD algorithms. The paper also provides a critical survey of existing supervised descriptive rule discovery visualization methods.

MSC:

68T05 Learning and adaptive systems in artificial intelligence

Cited in **15** Documents

Keywords:

descriptive rules; rule learning; contrast set mining; emerging patterns; subgroup discovery

Full Text: [Link](#)