

Knowledge discovery in real-world databases requires a broad scope of techniques and forms of knowledge. Both the knowledge and the applied methods should fit the discovery tasks and should adapt to knowledge hidden in the data. PKDD2001 Discovery Challenge will encourage a collaborative research effort, a broad and unified view of knowledge and methods of discovery, and emphasis on business problems and solutions to those problems. Unlike the competitive character of KDD Cups, we emphasize the collaborative nature of the knowledge discovery process.

The Discovery Challenge is organized around the Thrombosis data. This data set was used also in the past challenges in 1999 and 2000. One reason for this is rather pragmatic: it is difficult to get real world data set available. The second reason is, that the results obtained so far still have a preliminary character and the data experts are interested in further analyses. Because of collaborative paradigm of the challenge, the past results were available to all participants.

A session at the PKDD2001 conference has been organized to summarize the main findings and problems generated by the Challenge. 5 papers will be presented at this session. In the proceedings, the papers are ordered alphabetically according to the first author.

We wish to express our thanks to the data providers, the authors, as well as to all active Discovery Challenge participants.

We wish especially acknowledge the important contribution of Jan Zytkow. Jan not only came with the idea of organizing discovery challenges at the PKDD conferences but also did a lot of work when preparing and cleaning the Thrombosis data.

Petr Berka,
Shusaku Tsumoto,
Katsuhiko Takabayashi,
Shishir Gupta