ALBERT-LUDWIGS-
UNIVERSITÄT FREIBURG

FAKULTÄT FÜR ANGEWANDTE WISSENSCHAFTEN

Oberseminar

Zeit: Mittwoch, 13. 12. 1999, 17.00 Uhr
Ort: Geb.051, Raum 03-026, Am Flughafen 17

Building blocks for an inductive database approach to binary data mining.

Prof. Jean-Franois Boulicaut

INSA de Lyon

Abstract

One of the obvious hot topic in data mining research the last five years has been frequent set discovery from binary data. It concerns the discovery of set of attributes from large binary tables such that these attributes are true within a same line often enough. It is then easy to derive rules that describe the data e.g., the popular association rules. The Rule Mining and Generalized Databases project has been lauched at INSA Lyon in 1998 and aims at improving the state of the art for real-life database mining. Our guideline comes from the inductive database framework that has been suggested by Imielinski and Mannila (CACM, special issue on knowledge discovery, november 1996). Within that framework, the whole knowledge discovery process can be done by means of queries and we study querying primitives and their evaluation as well. After a simple formalization of the inductive database concept, the lecture will emphasize our last results. It concerns (a) the use of a multidimensional data manipulation language (a simple extension of datalog) at the different steps of a rule mining process and (b), the effective mining of concise representations of binary data, namely frequent sets and frequent closures. L. Lakhal and colleagues designed the close algorithm as an alternative to Apriori-like algorithms in difficult cases (ICDT'99). We introduce this algorithm: frequent closures can serve as generators for all the frequent sets but, in the case of correlated data, not only close remains tractable but also it provides a more concise representation of the data without any loss of information w.r.t. frequent sets. We discuss the experimental evidences we got from several experimentations and point out the perspectives of that technique.

Gastgeber: Prof. Dr. Luc De Raedt

Alle Interessenten sind recht herzlich eingeladen.

Organisation : Prof. Dr. B. Becker / F. Schmiedle

Weitere Informationen im WWW unter http://ira.informatik.uni-freiburg.de/FREAK/kolloquium oder von F.Schmiedle (Tel: 0761/203-8144; Email: schmiedl@informatik.uni-freiburg.de).