

Dortmund, 12. März 2003

EINLADUNG

Im Rahmen des Kolloquiums des Fachbereichs Informatik wird

Professor Abraham Bernstein, PhD
Institut für Informatik, Universität Zürich

einen Vortrag halten mit dem Titel

**Lunch is cheaper when you know a good recipe: towards a process oriented view of
KDD**

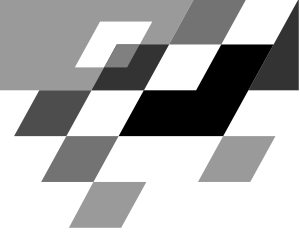
ORT: GB IV, H.112
ZEIT: Dienstag, 25. März 2003, 14:00 Uhr

ZUSAMMENFASSUNG:

Even though the knowledge discovery process (KD-process) has repeatedly been mentioned as being central to data mining (Fayyad, Piatetsky-Shapiro, & Smyth, 1996; Kerber, Beck, Anand, & Smart, 1998) few projects have directly focused on (investigating and) supporting the KD-process. Using examples from two ongoing projects this talk argues that using an appropriate data mining process is oftentimes more important than finding a good induction algorithm to support a call for more research about KD-processes. The first example introduces an Intelligent Discovery Assistant (IDA) to help novice and expert users explore the design space of KD-processes. Concretely, the IDA provides users with (i) systematic enumerations of valid KD processes, in order that important, potentially fruitful options are not overlooked, and (ii) effective rankings of these valid processes by different criteria, to facilitate the choice of KD processes to execute. Furthermore, it can serve as an important tool for knowledge sharing among a team of data miners. A case study exemplifies, how the process is oftentimes more important than the choice of an induction algorithm. The second example presents a case study about company affiliation identification (such as industry or sector) based on news stories. It presents a variety of KD processes to mine the source data basing on induction algorithms such as inductive logic programming, text mining, relational feature construction, and a novel relational vector space approach (in analogy to



ZU DIESEM VORTRAG LADEN HERZLICH EIN
DIE DOZENTEN DES FACHBEREICHES INFORMATIK



the vector space models used in information retrieval). Again the study seems to indicate that the largest differences in performance arise when varying the process rather than the induction algorithm.

Bio

Abraham Bernstein, heute Außerordentlicher Professor am Institut für Informatik an der Universität Zürich, studierte von 1988 bis 1994 an der ETH Zürich Informatik. Anschliessend an sein Studium arbeitete er zunächst für die Schweizerische Bankgesellschaft (UBS), zunächst am UBILAB, dem damaligen UBS-Forschungslabor für Informatik, und dann als Projektleiter für Informatikprojekte. Nach einer längeren Forschungstätigkeit am Center for Coordination Science des Massachusetts Institute of Technology (MIT), unter anderem als einer der Hauptmitarbeiter an der Forschung und Entwicklung des Process Handbook's, promovierte (Ph.D.) er im Jahre 2000 am MIT in Management mit Vertiefung in Information Technologies. Von 2000 bis 2002 war er als Assistenzprofessor für Informationssysteme an der Stern School of Business der New York University tätig. Abraham Bernsteins Forschungsinteressen beinhalten die computer-basierte Unterstützung von Prozessen, Prozessmodellierung, Computergestützte Gruppenarbeit (CSCW), Agentenbasierte Systeme, Data-Mining, das Semantic Web, und mobile Computertechnologien wie Wireless- und Ubiquitous-computing.

References

Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). The KDD Process for Extracting Usefull Knowledge from Volumes of Data. *Communication of the ACM*, 39(11), p. 27 - 34. Kerber, R., Beck, H., Anand, T., & Smart, B. (1998, August, 27-31, 1988). Active Templates: Comprehensive Support for the Knowledge Discovery Process. Paper presented at the Fourth International Conference on Knowledge Discovery and Data Mining, New York, NY.



ZU DIESEM VORTRAG LADEN HERZLICH EIN
DIE DOZENTEN DES FACHBEREICHES INFORMATIK