

EINLADUNG

Zeit: Mittwoch 14. September 2011, 15 Uhr

Ort: Raum 5056, Ahornstr. 55

Referent: Dipl.-Inform. Philipp Kranen
Lehrstuhl für Informatik 9, RWTH Aachen

Titel: Anytime Algorithms for Stream Data Mining

Abstract:

Data is collected and stored everywhere, be it images or audio files on private computers, customer data in traditional or electronic businesses, performance or control data in production sites, web traffic and click streams at internet providers, statistical data at government agencies, sensor measurements in scientific experimentation, surveillance data, etc. There are countless examples, and the amount of data is tremendous. Data mining is the process of finding useful and previously unknown patterns in data. In the examples listed above, data mining can be used for automated recommendation of audio files, business analysis and target marketing, or performance optimization and hazard warnings. While early mining algorithms only considered static data sets, research and practice in data mining must nowadays deal with continuous, possible infinite streams of data, which are prevalent in most real world applications and scenarios.

Anytime algorithms constitute a special type of algorithm that is well suited to work on data streams. They inherit their name from their ability to provide a result after any amount of processing time. The amount of time available is not known to the algorithm in advance: anytime algorithms quickly compute an initial result and strive to improve it as long as time remains. When interrupted they deliver the best result obtained until that point in time.

In this thesis anytime classification is studied in depth for the Bayesian approach. New algorithmic solutions for anytime classification are developed and evaluated in extensive experimentation. The first anytime stream clustering algorithm is proposed, and an application to anytime outlier detection is presented. In addition to the algorithmic contributions, new meta approaches are described that significantly widen the area of applications for anytime algorithms. The solutions and results of this thesis contribute to the state of the art in anytime algorithms and stream data mining research.

Es laden ein: Die Dozenten der Informatik