The new

Collaborative Research Center SFB 876 at the TU Dortmund University
Providing Information by Resource-Constrained Data Analysis

offers the following positions and opportunities:

Data Mining: 12 Postgraduate / Junior Researcher Positions, TU Dortmund University, SFB 876 with excellence in data mining, statistical learning, graphical models, parallel and distributed data mining, analysis of spectroscopy data, image analysis, high-throughput data analysis, or algorithm engineering

Embedded Systems: 6 Postgraduate / Junior Researcher Positions, TU Dortmund University, SFB 876 with excellence in embedded systems, operating systems of ubiquitous systems, communication networks, energy-aware algorithms

Information Systems: 2 Postgraduate / Junior Researcher Positions, TU Dortmund University, SFB 876 with excellence in Information engineering, multi-agent systems, privacy

Fellowship: 2 SFB 876 PhD or Postdoc Fellowships for 2 to 6 months

The collaborative research center SFB876 brings together data mining and embedded systems and enhances their methods such that information from distributed, dynamic masses of data becomes available anytime and anywhere. The research center SFB 876 is structured into 12 interdisciplinary projects. Positions for PhD students are now available for all of these projects. For further information about the projects see:

`http://www-ai.cs.uni-dortmund.de/auto?self=$fz10eqyn`

Conditions of Employment for Postgraduates

Salary is according to the German Federal pay scale (TV-L 13, approx. EUR 34,000 p.a.). The positions are available from January 2011 until the end of 2012 with the possibility of extension. Application deadline: 02/02/2011 or until a suitable candidate is found.

Women are particularly encouraged to apply. In case of equal personal aptitudes and qualification, disabled persons will be given priority. If applicable, the successful applicants are expected to work on obtaining a doctoral degree in the scope of this position.

How to Apply & What to Do in Case of Questions

Please address questions about the position and send your application in electronic form (PDF) by email to the project leader of the respective project and/or the speaker of the SFB876,
Prof. Dr. Katharina Morik (katharina.morik@tu-dortmund.de)

Please, indicate in the subject line of the mail: „*SFB876 application*" and the project identifier.

Postgraduates for Data Mining

Project A1 *Data Mining for Ubiquitous System Software*: Finding usage patterns in system log data of ubiquitous systems will be tackled in two ways. First, data are aggregated by streaming algorithms at the ubiquitous system, finding patterns or motifs that are well suited for further learning. Second, the log data are analysed by graphical models (e.g., CRF, SVMstruct), regarding special kernel functions. These algorithms are parallelized for CUDA implementation. Finally, distributed clustering algorithms are investigated in order to find groups of usage patterns.
Contact: **katharina.morik@tu-dortmund.de**

Project A2 *Algorithmic Aspects of Learning Methods in Embedded Systems*: Algorithmic foundations of learning methods in resource-constrained systems are investigated with respect to memory models. In particular, clustering algorithms, graph algorithms, distributed algorithms, and data streaming algorithms are investigated with respect to energy consumption. Developing the theoretical basis for resource-aware learning is the focus of A2. Contact: **christian.sohler@tu-dortmund.de**; **jan.vahrenhold@tu-dortmund.de**

Project A3 *Methods for the Efficient Resource Utilization in Machine Learning Algorithms*: R programs are analyzed w.r.t. their resource consumption and a well-tailored compilation strategy will be developed. For diverse statistical learning algorithms the trade-off between quality of the learned model and recourse consumption will be determined. In particular, problems of model selection, feature selection will be investigated. Contact: **joerg.rahnenfuehrer@tu-dortmund.de**

Project TB1 *Analysis of Spectrometry Data with Restricted Resources*: The analysis of spectroscopy data from a breath sensor for the timely disease diagnosis requires an online method for feature extraction, alignment methods, and prediction of spectra. High-throughput data analysis, bioinformatics, and online algorithms are the fields that contribute to developing smart wear including a breath sensor. Contact: **sven.rahmann@tu-dortmund.de**

Project B2 *Resource-optimizing time analysis of artifactious image sequences for the detection of nano objects*: Plasmon assisted microscopy of nano objects (PAMONO) is a new sensor measuring extremely little objects. The PAMONO images are to be analysed in order to detect and classify viruses. The images are suffering from noise. Cleaning the data, comparing image analysis algorithms regarding their recognition capabilities and their respective resource consumptions are the tasks of data mining in this project. Contact: **heinrich.mueller@tu-dortmund.de**; **alexander.zybin@isas.de**

Project C1 *Feature Selection in High Dimensional Data for Risk Prognosis in Oncology*: Ensemble methods are to be investigated, in how much they increase the stability of feature selection in high dimensional data where much less examples are given than features. Sets of exons (several exons make up for a gene) are to be clustered w.r.t. the oncology prognosis. Transfer learning will be evaluated concerning example sets stemming from different methods and collections of patients. Co-training will combine and fuse divers studies on exons or genes for neuroblastoma patients. Contact: **katharina.morik@tu-dortmund.de**

Project C2 *Multi-level statistical analysis of high-frequency spatio-temporal process data*: High-frequency high-dimensional process data are typically analyzed in several steps. Interesting data need to be identified and selected in real time for further analysis because of limited storage capacities. By means of the challenging example of detection and classification of astrophysical particles appearing as transient patterns in spatio-temporal data, we want to construct robust generalizable techniques for reduction, classification and modeling of such data, as well as strategies for the suitable combination of these analysis steps. Contact: **roland.fried@tu-dortmund.de**; **wolfgang.rhode@tu-dortmund.de**

Project C3 *Regression approaches for large-scale high-dimensional data*: The scalability of modern regression approaches is often stretched to its limits by a large number of observations and/or variables. This aggravates their use in embedded systems. The goal of this project is therefore the development of highly efficient regression methods. Algorithms are to be developed that reduce the number of observations using, e.g., random linear projections and sampling (streaming algorithms), or the dimensionality of the underlying, possibly Bayesian, model classes imposing structural constraints, e.g., monotonicity. Contact: **katja.ickstadt@tu-dortmund.de**; **christian.sohler@tu-dortmund.de**

Postgraduates for Embedded Systems

Project A1 *Data Mining for Ubiquitous System Software*: Instrumenting the operating system of ubiquitous systems for data gathering is difficult, since logging the data must not disturb the regular use of the system. The instrumentation must be platform independent. New methods based on Dynamic Aspect C++ are to be developed and optimized regarding resource consumption. Applying learned models to the operating systems requires a generic framework so that efficiency is actually enhance by them. Contact: **olaf.spinczyk@tu-dortmund.de**

Project A3 *Methods for the Efficient Resource Utilization in Machine Learning Algorithms*: This project builds a bridge between learning algorithms in terms of R programs and resource efficiency. Based on an internal representation of R, optimizations of the compiler are to be developed that take into account memory allocations. The optimized internal representation is then transformed until a stanrad compiler and C++ libraries produce resource-efficient code.
Contact: **peter.marwedel@tu-dortmund.de**

Project A4 *Resource Efficient and Distributed Platforms for Integrative Aata Analysis*: This project explores technical and conceptual foundations for distributed execution platforms that are self-descriptive regarding their provision and consumption of resources. In particular, the description of resources like memory, power consumption, bandwidth, or runtime constraints is of special interest. Due to the platform's ability to describe itself, applications have a valid basis for resource-aware decisions and hence are enabled to optimize their resource requests. The major challenge of this approach is to deal with the vast complexity of embedded platforms and dynamic interactions between the resource types envisaged. Contact: **olaf.spinczyk@tu-dortmund.de**

Project B2 *Resource-optimizing Time Analysis of Artifactious Image Sequences for the Detection of Nano Objects*: The platform from project A4 is enhanced. Here, FPGA approaches and implementations for GPGPU are developed. In order to gain independence from various multiprocessor architectures, a high-level synthesis and multi-objective optimization of hardware-software partitionings are to be investigated. Contact: **peter.marwedel@tu-dortmund.de**

Project B4 *Analysis and Communication for Dynamic Traffic Prognosis*: This project researches innovative, high precision and real time prognosis algorithms to predict object stream behaviour. The quality of the prognosis depends on the quantity, quality and reliability plus synchronicity of the available information on representative objects in the stream. In consideration of the avaiability of the multitude of sensor data sources as well as resource starved embedded systems, this sub-project also explores novel data aggregation, fusion and communication methods with respect to the requirements set by the prognosis, and evaluates the expected benefits.
Contact: **christian.wietfeld@tu-dortmund.de**; **schreckenberg@ptt.uni-due.de**

Postgraduates for Information Engineering

Project A5 *Exchange and Fusion of Information under Availability and Confidentiality Requirements in Multi-Agent Systems*: Achieving both availability and confidentiality of information implies the development and employment of concepts and algorithms that handle conflicting security issues of actors and persons. This should be possible even in situations with limited technical infrastructure in terms of time, space, and communication. This project aims at extending data base techniques to multi-agent systems oriented towards the approach of J.Y. Halpern and K.R. O'Neill (2005, 2008), in particular for the purpose of data mining in embedded systems. Contact: **joachim.biskup@tu-dortmund.de**; **gabriele.kern-isberner@tu-dortmund.de**

SFB/TR 8 Postdoc Fellowships for 2 to 6 months

The research center SFB 876 offers fellowships for visiting postgraduate or postdoctoral researchers for a lump sum of EUR 3,000 per month. Candidates from all over the world and with excellence in computer science particularly in data analysis or embedded systems will be considered, if they demonstrate potential to engage in innovative research together with SFB 876 researchers and students. Applicants need to propose a research program for two to six months, linked to one or more projects of the SFB 876. Contact: **katharina.morik@tu-dortmund.de**; **wolfgang.rhode@tu-dortmund.de**