

# Preliminary Program of IDA 2003

(August 20, 2003)

---

## Thursday, August 28th, 2003

---

9:00 – 12:00 **Tutorials**

(in parallel, tutorials are free for registered attendees)

12:00 – 13:30 Lunch Break

13:30 – 14:00 **Opening Address, Room 101**

*Prof. D. Lenzen, President of the Free University of Berlin*

*Prof. H. Bester, Dean of the Faculty of Economics*

*Prof. H.-J. Lenz, General Chairman of IDA 2003*

14:00 – 15:30 **Machine Learning 1**

Pruning for Monotone Classification Trees

*A. Feelders and M. Pardoel*

Regularized Learning with Flexible Constraints

*E. Hüllermeier*

Learning to Answer Emails

*A. Lüneburg, M. Kockelkorn, and T. Scheffer*

15:30 – 16:00 Coffee Break

16:00 – 17:30 **Probability and Topology**

Learning Dynamic Bayesian Networks from Multivariate Time Series with Changing Dependencies

*A. Tucker and X. Liu*

Topology and Intelligent Data Analysis

*V. Robins, J. Abernethy, N. Rooney, and L. Bradley*

Coherent Conditional Probability as a Measure of Information of the Relevant Conditioning Events

*G. Coletti and R. Scozzafava*

19:30 – **Dinner on the Reichstag**

Departure from conference site: 18:30 hours.

Please have a passport or an id card ready at the entrance.

---

## Friday, August 29th, 2003

---

9:00 – 10:00 **Keynote Talk 1**

A Virtual Cup-Tester for Coffee Quality Assessment

*G. della Riccia*

10:00 – 10:30 Coffee Break

10:30 – 12:00 **Classification and Pattern Recognition**

Learning Linear Classifiers Sensitive to Example Dependent and Noisy Costs

*P. Geibel and F. Wyszotki*

An Effective Associative Memory for Pattern Recognition

*B. Kryzhanovsky, L. Litinskii, and A. Fonarev*

Similarity Based Classification

*A.E. Bernal, K. Hovsepian, T. Karadeniz, and J.-L. Lassez*

12:00 – 13:00 Lunch Break

13:00 – 14:30 **Machine Learning 2**

A Semi-supervised Method for Learning the Structure of Robot Environment Interaction

*A. Grossmann, M. Wendt, and J. Wyatt*

Using Domain Specific Knowledge for Automated Modeling

*L. Todorovski and S. Dzeroski*

Resolving Rule Conflicts with Double Induction

*T. Lindgren and H. Boström*

14:30 – 14:45 Coffee Break

14:45 – 16:15 **Clustering**

What is Fuzzy about Fuzzy Clustering?

Understanding and Improving the Concept of the Fuzzifier

*F. Klawonn and F. Höppner*

A Mixed Model Approach for Binned Data Clustering

*A. Same, C. Ambroise, and G. Govaert*

16:15 – 16:30 Coffee Break

16:30 – 17:30 **Poster Spotlight Presentations**

17:30 – 19:00 **Poster Presentations**

**Machine Learning**

A Novel Partial-Memory Learning Algorithm based on Grey Relational Structure

*C.-C. Huang and H.-M. Lee*

**Classification and Pattern Recognition**

Numerical Attributes in Decision Trees: A Hierarchical Approach

*F. Berzal, J.-C. Cubero, N. Marin, and D. Sanchez*

Similarity-based Neural Networks for Applications on Symbol Strings

*I. Fischer*

Combining Pairwise Classifiers with Stacking

*P. Savicky and J. Fuernkranz*

**Clustering**

Fuzzy Clustering based Segmentation of Time-series

*J. Abonyi, B. Feil, S. Nemeth, and P. Arva*

An Iterated Local Search Approach for Minimum Sum-Of-Squares Clustering

*P. Merz*

Data Clustering in Tolerance Space

*C.-H. Tzeng and F.-S. Sun*

**Probability and Topology**

Very Predictive Ngrams for Space-Limited Probabilistic Models

*P. Cohen and C. Sutton*

Interval Estimation Naïve Bayes

*V. Robles, P. Larrañaga, J.M. Peña, E. Menasalvas, and M.S. Pérez*

**Modeling**

A Multiagent-Based Constructive Approach for Feedforward Neural Networks

*C.A.M. Lima, A.L.V. Coelho, and F.J. Von Zuben*

Evolutionary System Identification via Descriptive Takagi Sugeno Fuzzy Systems

*I. Renners and A. Grauel*

Minimum Message Length Criterion for Second-order Polynomial Model Selection Applied to Tropical Cyclone Intensity Forecasting

*G.W. Rumantir and C.S. Wallace*

On the Use of the GTM Algorithm for Mode Detection

*S. Vegas-Azcarate and J. Muruzabal*

#### **(Data) Preprocessing**

A Logical Formalisation of the Fellegi-Holt Method of Data Cleaning

*A. Boskovitz, R. Gore, and M. Hegland*

Compression Technique Preserving Correlations of a Multivariate Temporal Sequence

*A. Chouakria-Douzal*

Condensed Representations in Presence of Missing Values

*F. Rioult and B. Cremilleux*

#### **Applications**

Genome-wide Prokaryotic Promoter Recognition based on Sequence Alignment Kernel

*L. Gordon, A. Chervonenkis, A. Gammerman, I. Shahmuradov, and V. Solovyev*

Towards Automated Electrocardiac Map Interpretation: An Intelligent Contouring Tool based on Spatial Aggregation

*L. Ironi and S. Tentoni*

Gaussian Mixture Density Estimation Applied to Microarray Data

*C. Steinhoff, T. Müller, U. Nuber, and M. Vingron*

19:00 – **Barbeque**

---

### **Saturday, August 30th, 2003**

---

9:00 – 10:00 **Keynote Talk 2**

Distributed Learning for the Analysis of Extreme Data Sets

*L.O. Hall*

10:00 – 10:30 Coffee Break

10:30 – 12:00 **Applications (Bioinformatics)**

Selective Sampling with a Hierarchical Latent Variable Model

*H. Mamitsuka*

Obtaining Quality Microarray Data via Image Reconstruction

*P. O'Neill, G.D. Magoulas, and X. Liu*

Large Scale Mining of Molecular Fragments with Wildcards

*H. Hofer, C. Borgelt, and M. Berthold*

12:00 – 13:00 Lunch Break

13:00 – 14:00 **Poster Spotlight Presentations**

14:00 – 15:30 **Poster Presentations**

#### **Machine Learning**

Constructing Hierarchical Rule Systems

*T.R. Gabriel and M.R. Berthold*

Text Categorization Using Hybrid Multiple Model Schemes

*I.-C. Kim, S.-H. Myoung*

#### **Classification and Pattern Recognition**

APRIORI-SD: Adapting Association Rule Learning to Subgroup Discovery

*B. Kavsek, N. Lavrac, V. Jovanoski*

Solving Classification Problems using Infix Form Genetic Programming

*M. Oltean and C. Grosan*

### **Clustering**

Refined Shared Nearest Neighbors Graph for Combining Multiple Data Clusterings

*H. Ayad and M. Kamel*

Clustering Mobile Trajectories for Resource Allocation in Mobile Environments

*D. Katsaros, A. Nanopoulos, M. Karakaya, G. Yavas, U. Ulusoy, and Y. Manolopoulos*

Fuzzy Clustering of Short Time Series and Unevenly Distributed Sampling Points

*C. Moeller-Levet, F. Klawonn, K.-H. Cho, and O. Wolkenhauer*

Combining and Comparing Cluster Methods in a Receptor Database

*E.V. Samsonova, T. Baeck, M.W. Beukers, A.P. Ijzerman, and J.N. Kok*

### **Probability and Topology**

Mining Networks and Central Entities in Digital Libraries: A Graph Theoretic Approach applied to Co-Author Networks

*P. Mutschke*

### **Modeling**

Regularization Methods for Additive Models

*M. Avalos, Y. Grandvalet, and C. Ambroise*

Automated Detection of Influenza Epidemics with Hidden Markov Models

*T. Rath, M. Carreras, and P. Sebastiani*

Guided Incremental Construction of Belief Networks

*C. Sutton, B. Burns, C. Morrison, and P. Cohen*

Distributed Regression For Heterogeneous Data Sets

*Y. Xing, M.G. Madden, J. Duggan, and G. Lyons*

### **(Data) Preprocessing**

Measures of Rule Quality for Feature Selection in Text Categorization

*E. Montañés, J. Fernández, I. Díaz, E.F. Combarro, and J. Ranilla*

Genetic Approach to Constructive Induction Based on Non-algebraic Feature Representation

*L. Shafti and E. Perez*

Active Feature Selection based on a Very Limited Number of Entities

*Y. Sinohara and T. Miura*

### **Applications**

Classification of Protein Localisation Patterns via Supervised Neural Network Learning

*A.D. Anastasiadis, G.D. Magoulas, and X. Liu*

Applying Intelligent Data Analysis to Coupling Relationships in Object-oriented Software

*S. Counsell, X. Liu, R. Najjar, S. Swift, and A. Tucker*

The Smaller the Better: Comparison of Two Approaches for Sales Rate Prediction

*M. Lauer, M. Riedmiller, T. Ragg, W. Baum, and M. Wigbers*

17:30 –

### **Boat Trip**

Departure from conference site: 16:30 hours