ECML/PKDD-2002 PROGRAM SCHEDULE

MONDAY 19 August 2002

9:00 **Tutorial T1: Text Mining and Internet Content Filtering** — **Sali 1**
José María Gómez Hidalgo

Tutorial T2: Formal Concept Analysis — **Sali 5**
Gerd Stumme

Workshop WS1: IDDM — Integration and Collaboration Aspects of Data Mining, Decision Support and Meta-Learning — **Sali 13**
Organizers: Marko Bohanec, Dunja Mladenić, Nada Lavrač

Workshop WS2: VDM — Visual Data Mining — **Sali 6**
Organizers: Simeon J. Simoff, Monique Noirhomme-Fraiture, Michael H. Böhlen

Coffee will be served 10:30–11:00

13:00 **Lunch Break**

14:30 **Tutorial T3: Web Usage Mining for E-business Applications** — **Sali 1**
Myra Spiliopoulou, Bamshad Mobasher, Bettina Berendt

Tutorial T4: Inductive Databases and Constraint-based Mining — **Sali 5**
Jean-François Boulicaut, Luc De Raedt

Workshop WS1: IDDM — Integration and Collaboration Aspects of Data Mining, Decision Support and Meta-Learning — **Sali 13**
Program continues from morning

Workshop WS2: VDM — Visual Data Mining — **Sali 6**
Program continues from morning

Coffee will be served 16:00–16:30

18:00 **End of Tutorials and Workshops**

TUESDAY 20 August 2002

9:00 **Tutorial T5: An introduction to Quality Assessment in Data Mining** — **Sali 1**
Maria Halkidi, Michalis Vazirgiannis

Tutorial T6: Privacy, Security, and Data Mining — **Sali 12**
Chris Clifton

Workshop WS4: SWM — Semantic Web Mining — **Sali 13**
Organizers: Bettina Berendt, Andreas Hotho, Gerd Stumme

Workshop WS5: MOD — Mining Official Data — **Sali 6**
Organizers: Paula Brito, Donato Malerba

Workshop WS6: KDID — Knowledge Discovery in Inductive Databases — **Sali 10**
Organizers: Mika Klemettinen, Rosa Meo, Fosca Giannotti, Luc De Raedt

Workshop WSC: Discovery Challenge — **Auditorio XIII**
Organizers: Petr Berka, Jan Rauch, Shusaku Tsumoto

European Project Presentations Day — **Sali 5**
Organizers: Dunja Mladenić, Maarten van Someren

Coffee will be served 10:30–11:00

13:00 **Lunch Break**

14:30 **Workshop WS4: SWM — Semantic Web Mining** — **Sali 13**
Program continues from morning

Workshop WS5: MOD — Mining Official Data — **Sali 6**
Program continues from morning

Workshop WS6: KDID — Knowledge Discovery in Inductive Databases — **Sali 10**
Program continues from morning

Workshop WSC: Discovery Challenge — **Auditorio XIII**
Program continues from morning

European Project Presentations Day: Exhibition — **Lobby 2nd Floor**

16:00 **Coffee Break**

16:30 **ECML/PKDD-2002 Opening and Keynote Talk** — **Juhlasali**
Padhraic Smyth: *Learning with Mixture Models: Concepts and Applications*
Chair: Heikki Mannila

18:00 **Reception Hosted by the City of Helsinki** — **City Hall**
Address: Pohjoisesplanadi 11–13, 00170 Helsinki

WEDNESDAY 21 August 2002

9:00 **Keynote Talk** — **Sali 1**
Bernhard Schölkopf: *Learning with Kernels*
Chair: Tapio Elomaa

10:00 **KDNet Best ECML-2002 Paper Award Winner** — **Sali 1**
Ioannis Tsochantaridis, Thomas Hofmann: *Support Vector Machines for Polycategorical Classification*

10:30 **Coffee Break**

11:00 **Session A1: Ensemble Methods** — **Sali 1**
Chair: Luc De Raedt

- $\mathcal{E}$ Kwek, Nguyen: *iBoost: Boosting Using an Instance-Based Exponential Weighting Scheme*
- $\mathcal{E}$ Derbeko, El-Yaniv, Meir: *Variance Optimized Bagging*
- $\mathcal{P}$ Choki, Suzuki: *Iterative Data Squashing for Boosting Based on a Distribution-Sensitive Distance*
- $\mathcal{E}$ Fürnkranz: *Pairwise Classification as an Ensemble Technique*

Session A2: Text and Document Mining — **Sali 5**
Chair: Helena Ahonen-Myka

- $\mathcal{P}$ Forman: *Choose Your Words Carefully: An Empirical Study of Feature Selection Metrics for Text Classification*
- $\mathcal{P}$ Abe, Kawasoe, Asai, Arimura, Arikawa: *Optimized Substructure Discovery for Semi-Structured Data*
- $\mathcal{P}$ Winkler, Spiliopoulou: *Structuring Domain-Specific Text Archives by Deriving a Probabilistic XML DTD*
- $\mathcal{P}$ Kosala, Van den Bussche, Bruynooghe, Blockeel: *Information Extraction in Structured Documents using Tree Automata Induction*

Session A3: Semi-supervised and Unsupervised Learning — **Sali 13**
Chair: Katharina Morik (to be confirmed)

- $\mathcal{P}$ Giannotti, Gozzi, Manco: *Clustering Transactional Data*
- $\mathcal{E}$ Vittaut, Amini, Gallinari: *Learning Classification with Both Labeled and Unlabeled Data*
- $\mathcal{P}$ Ding, He, Zha, Simon: *Unsupervised Learning: Self-Aggregation in Scaled Principal Component Space*
- $\mathcal{E}$ Todorovski, Blockeel, Džeroski: *Ranking with Predictive Clustering Trees*

13:00 **Lunch Break**

14:30 **Session B1: Support Vector Machines** — **Sali 1**
Chair: Stephen Kwek

- $\mathcal{E}$ Engel, Mannor, Meir: *Sparse Online Greedy Support Vector Regression*
- $\mathcal{E}$ Martin: *On-line Support Vector Machine Regression*
- $\mathcal{P}$ Paaß, Leopold, Larson, Kindermann, Eickeler: *SVM Classification Using Sequences of Phonemes and Syllables*

Session B2: Relational Learning — **Sali 5**
Chair: Stan Matwin

- $\mathcal{E}$ Peña Castillo, Wrobel: *Macro-Operators in Multirelational Learning: A Search-Space Reduction Technique*
- $\mathcal{E}$ Ženko, Džeroski: *Stacking with an Extended Set of Meta-Level Attributes and MLR*
- $\mathcal{P}$ Knobbe, Siebes, Marseille: *Involving Aggregate Functions in Multi-Relational Search*

Session B3: Time Series and Collaborative Recommendation — **Sali 13**
Chair: Einoshin Suzuki

- $\mathcal{P}$ Berberidis, Vlahavas, Aref, Atallah, Elmagarmid: *On the Discovery of Weak Periodicities in Large Time Series*
- $\mathcal{P}$ Hirano, Tsumoto: *Multiscale Comparison of Temporal Patterns in Time-Series Medical Databases*
- $\mathcal{E}$ Kushmerick: *Robustness Analyses of Instance-Based Collaborative Recommendation*

16:00 **Coffee Break**

16:30 **Session C1: Biological and Medical Applications** — **Sali 1**
Chair: Maarten van Someren

- $\mathcal{P}$ Arnborg, Agartz, Hall, Jönsson, Sillén, Sedvall: *Data Mining in Schizophrenia Research — Preliminary Analysis*
- $\mathcal{P}$ Lidén, Asker, Boström: *Rule Induction for Classification of Gene Expression Array Data*
- $\mathcal{P}$ Tsumoto: *Mining Hiearchical Decision Rules from Clinical Databases using Rough Sets and Medical Diagnostic Model*

Session C2: Association Rules — **Sali 5**
Chair: Jean-François Boulicaut

- $\mathcal{P}$ Sese, Morishita: *Answering the Most Correlated N Association Rules Efficiently*
- $\mathcal{P}$ Coenen, Leng: *Finding Association Rules with Some Very Frequent Attributes*
- $\mathcal{P}$ Hüllermeier: *Association Rules for Expressing Gradual Dependencies*

Session C3: Sampling and Feature Selection — **Sali 13**
Chair: Johannes Fürnkranz

- $\mathcal{P}$ Klösgen, May: *Spatial Subgroup Mining Integrated in an Object-Relational Spatial Database*
- $\mathcal{P}$ Mamitsuka: *Iteratively Selecting Feature Subsets for Mining from High-Dimensional Databases*
- $\mathcal{P}$ Scheffer, Wrobel: *A Scalable Constant-Memory Sampling Algorithm for Pattern Discovery in Large Databases*

18:00 **End of Sessions C1 – C3**

19:30 **Conference Dinner** — **Restaurant Kulosaaren Casino**
Boats leave from the Market Square at 19:30
Address of the restaurant: Hopeasalmenpolku 1, 00570 Helsinki

THURSDAY 22 August 2002

9:00 **Keynote Talk** — **Sali 1**
Dan Roth: *Reasoning with Classifiers*
Chair: Hannu Toivonen

10:00 **KDNet Best PKDD-2002 Paper Award Winner** — **Sali 1**
Toon Calders, Bart Goethals: *Mining All Non-Derivable Frequent Itemsets*

10:30 **Coffee Break**

11:00 **Session D1: Data Mining** — **Sali 1**
Chair: Michael May

- $\mathcal{P}$ Bailey, Manoukian, Ramamohanarao: *Fast Algorithms for Mining Emerging Patterns*
- $\mathcal{P}$ Li, Wong: *Geography of Differences between Two Classes of Data*
- $\mathcal{P}$ Jaroszewicz, Simovici: *Support Approximations Using Bonferroni-Type Inequalities*
- $\mathcal{P}$ Kargupta, Sivakumar, Ghosh: *Dependency Detection in MobiMine and Random Matrices*

Session D2: Ensemble Methods — **Sali 5**
Chair: Stefan Wrobel

- $\mathcal{E}$ Eibl, Pfeiffer: *How to Make AdaBoost.M1 Work for Weak Base Classifiers by Changing Only One Line of the Code*
- $\mathcal{E}$ Nock, Lefaucheur: *A Robust Boosting Algorithm*
- $\mathcal{E}$ Thollard, Sebban, Ezequel: *Boosting Density Function Estimators*
- $\mathcal{P}$ Wall, Cunningham, Walsh: *Explaining Predictions from a Neural Network Ensemble One at a Time*

Session D3: Class Probability Estimation — **Sali 13**
Chair: Bernhard Pfahringer

- $\mathcal{E}$ Halck: *Using Hard Classifiers to Estimate Conditional Class Probabilities*
- $\mathcal{E}$ Kukar, Kononenko: *Reliable Classifications with Machine Learning*
- $\mathcal{E}$ Margineantu: *Class Probability Estimation and Cost-Sensitive Classification Decisions*
- $\mathcal{E}$ Banerjee, Peng: *Convergent Gradient Ascent in General-Sum Games*

13:00 **Lunch Break**

14:30 **Session E1: Reinforcement Learning** — **Sali 1**
Chair: Joost N. Kok

- $\mathcal{E}$ Menache, Mannor, Shimkin: *Q-Cut — Dynamic Discovery of Sub-Goals in Reinforcement Learning*
- $\mathcal{E}$ Preux: *Propagation of Q-values in Tabular TD(λ)*

Session E2: Web Mining — **Sali 5**
Chair: Myra Spiliopoulou

- $\mathcal{P}$ Park, Palaniswami, Ramamohanarao: *A Novel Web Text Mining Method Using the Discrete Cosine Transform*
- $\mathcal{P}$ Kemp, Ramamohanarao: *Long-Term Learning for Web Search Engines*

Session E3: Relational Learning — **Sali 13**
Chair: Arno Siebes

- $\mathcal{E}$ Hoche, Wrobel: *Scaling Boosting by Margin-Based Inclusion of Features and Relations*
- $\mathcal{E}$ Morik, Rüping: *A Multistrategy Approach to the Classification of Phases in Business Cycles*

15:30 **Coffee Break**

16:00 **Session F1: Decision Trees** — **Sali 1**
Chair: Hendrik Blockeel

- $\mathcal{E}$ Holmes, Pfahringer, Kirkby, Frank, Hall: *Multiclass Alternating Decision Trees*
- $\mathcal{E}$ Hüllermeier: *Possibilistic Induction in Decision-Tree Learning*

Session F2: Markov Techniques — **Sali 5**
Chair: Tobias Scheffer

- $\mathcal{E}$ Kermorvant, Dupont: *Improved Smoothing for Probabilistic Suffix Trees Seen as Variable Order Markov Chains*
- $\mathcal{E}$ Ratitch, Precup: *Characterizing Markov Decision Processes*

Session F3: Association Rules — **Sali 13**
Chair: Kotagiri Ramamohanarao (to be confirmed)

- $\mathcal{P}$ Jeudy, Boulicaut: *Using Condensed Representations for Interactive Association Rule Mining*
- $\mathcal{P}$ Veloso, Gusmão, Meira, Carvalho, Parthasarathy, Zaki: *Efficiently Mining Approximate Models of Associations in Evolving Databases*

17:00 **ECML/PKDD Business Meeting** — **Sali 5**

18:00 **Reception Hosted by the University of Helsinki** — **Lehtisali**

FRIDAY 23 August 2002

9:00 **Keynote Talk** — **Sali 1**
Erkki Oja: *Finding Hidden Factors Using Independent Component Analysis*
Chair: Heikki Mannila

10:00 **Coffee Break**

10:30 **Session G1: Aspects of Data Mining** — **Sali 1**
Chair: Fosca Giannotti (to be confirmed)

- $\mathcal{P}$ Angiulli, Pizzuti: *Fast Outlier Detection in High Dimensional Spaces*
- $\mathcal{P}$ Felty, Matwin: *Privacy-Oriented Data Mining by Proof Checking*
- $\mathcal{P}$ Gamberger, Lavrač: *Generating Actionable Knowledge by Expert-Guided Subgroup Discovery*
- $\mathcal{P}$ Koyutürk, Grama, Ramakrishnan: *Algebraic Techniques for Analysis of Large Discrete-Valued Datasets*

Session G2: Foundations of Learning — **Sali 5**
Chair: Sašo Džeroski

- $\mathcal{E}$ Rückert, Kramer, De Raedt: *Phase Transitions and Stochastic Local Search in k-Term DNF Learning*
- $\mathcal{E}$ Harris: *Evidence that Incremental Delta-Bar-Delta is an Attribute-Efficient Linear Learner*
- $\mathcal{E}$ Papadopoulos, Proedrou, Vovk, Gammerman: *Inductive Confidence Machines for Regression*
- $\mathcal{E}$ Proedrou, Nouretdinov, Vovk, Gammerman: *Transductive Confidence Machines for Pattern Recognition*

Session G3: Supervised Learning — **Sali 13**
Chair: Donato Malerba

- $\mathcal{E}$ Dai, Li, Tu: *An Empirical Study of Encoding Schemes and Search Strategies in Discovering Causal Networks*
- $\mathcal{E}$ Góra, Wojna: *RIONA: A Classifier Combining Rule Induction and k-NN Method with Automated Selection of Optimal Neighbourhood*
- $\mathcal{P}$ Zighed, Lallich, Muhlenbach: *Separability Index in Supervised Learning*
- $\mathcal{P}$ Brain, Webb: *The Need for Low Bias Algorithms in Classification Learning from Large Data Sets*

12:30 **Lunch Break**

14:00 **Session H1: Unsupervised Learning** — **Sali 1**
Chair: Stefan Kramer

- $\mathcal{E}$ Ludl, Widmer: *Towards a Simple Clustering Criterion Based on Minimum Length Encoding*
- $\mathcal{E}$ Buntine: *Variational Extensions to EM and Multinomial PCA*
- $\mathcal{E}$ Sinkkonen, Kaski, Nikkilä: *Discriminative Clustering: Optimal Contingency Tables by Learning Metrics*

Session H2: Applications of Learning — **Sali 5**
Chair: Michèle Sebag

- $\mathcal{E}$ Carreras, Màrquez, Punyakanok, Roth: *Learning and Inference for Clause Identification*
- $\mathcal{E}$ Klink, Hust, Junker, Dengel: *Collaborative Learning of Term-Based Concepts for Automatic Query Expansion*
- $\mathcal{E}$ Kråkenes, Halck: *Learning to Play a Highly Complex Game from Human Expert Games*

Session H3: Learning Approaches — **Sali 13**
Chair: Richard Nock

- $\mathcal{E}$ Ontañón, Plaza: *Case Exchange Strategies in Multiagent Learning*
- $\mathcal{E}$ Bay, Shapiro, Langley: *Revising Engineering Models: Combining Computational Discovery with Knowledge*
- $\mathcal{P}$ Joshi, Agarwal, Kumar: *Predicting Rare Classes: Comparing Two-Phase Rule Induction to Cost-Sensitive Boosting*
- $\mathcal{P}$ Domeniconi, Perng, Vilalta, Ma: *A Classification Approach for Prediction of Target Events in Temporal Sequences*

16:00 **End of Sessions H1 – H3**

Key

$\mathcal{E}$ ECML paper

$\mathcal{P}$ PKDD paper

Sali 1, **Sali 5**, **Sali 6**, **Sali 10**, **Sali 12**, **Sali 13** — lecture halls or lecture rooms in the New Part of the Main Building of the University of Helsinki

Auditorio XIII a lecture room in the Old Part of the Main Building

Juhlasali the Main Hall in the Old Part of the Main Building

Lehtisali the Reading Room in the Old Part of the Main Building

New Part Fabianinkatu side of the Main Building

Old Part Unioninkatu side of the Main Building