

Table of Contents – Part II

Privacy and Security

<i>AUDIO: An Integrity <u>A</u>uditing Framework of Outlier-Mining-as-a-Service Systems</i>	1
<i>Ruilin Liu, Hui (Wendy) Wang, Anna Monreale, Dino Pedreschi, Fosca Giannotti, and Wenge Guo</i>	
Differentially Private Projected Histograms: Construction and Use for Prediction	19
<i>Staal A. Vinterbo</i>	
Fairness-Aware Classifier with Prejudice Remover Regularizer	35
<i>Toshihiro Kamishima, Shotaro Akaho, Hideki Asoh, and Jun Sakuma</i>	

Rankings and Recommendations

A Live Comparison of Methods for Personalized Article Recommendation at Forbes.com	51
<i>Evan Kirshenbaum, George Forman, and Michael Dugan</i>	
Fast ALS-Based Tensor Factorization for Context-Aware Recommendation from Implicit Feedback	67
<i>Balázs Hidasi and Domonkos Tikk</i>	
Probability Estimation for Multi-class Classification Based on Label Ranking	83
<i>Weiwei Cheng and Eyke Hüllermeier</i>	

Reinforcement Learning and Planning

Adaptive Planning for Markov Decision Processes with Uncertain Transition Models via Incremental Feature Dependency Discovery	99
<i>N. Kemal Ure, Alborz Geramifard, Girish Chowdhary, and Jonathan P. How</i>	
APRIL: Active Preference Learning-Based Reinforcement Learning	116
<i>Riad Akrouf, Marc Schoenauer, and Michèle Sebag</i>	
Autonomous Data-Driven Decision-Making in Smart Electricity Markets	132
<i>Markus Peters, Wolfgang Ketter, Maytal Saar-Tsechansky, and John Collins</i>	

Bayesian Nonparametric Inverse Reinforcement Learning	148
<i>Bernard Michini and Jonathan P. How</i>	
Bootstrapping Monte Carlo Tree Search with an Imperfect Heuristic	164
<i>Truong-Huy Dinh Nguyen, Wee-Sun Lee, and Tze-Yun Leong</i>	
Fast Reinforcement Learning with Large Action Sets Using Error-Correcting Output Codes for MDP Factorization	180
<i>Gabriel Dulac-Arnold, Ludovic Denoyer, Philippe Preux, and Patrick Gallinari</i>	
Learning Policies for Battery Usage Optimization in Electric Vehicles . . .	195
<i>Stefano Ermon, Yexiang Xue, Carla Gomes, and Bart Selman</i>	
Policy Iteration Based on a Larned Transition Model	211
<i>Vivek Ramavajjala and Charles Elkan</i>	
Structured Apprenticeship Learning	227
<i>Abdeslam Boularias, Oliver Krömer, and Jan Peters</i>	

Rule Mining and Subgroup Discovery

A Bayesian Approach for Classification Rule Mining in Quantitative Databases	243
<i>Dominique Gay and Marc Boullé</i>	
A Bayesian Scoring Technique for Mining Predictive and Non-Spurious Rules	260
<i>Iyad Batal, Gregory Cooper, and Milos Hauskrecht</i>	
Generic Pattern Trees for Exhaustive Exceptional Model Mining	277
<i>Florian Lemmerich, Martin Becker, and Martin Atzmueller</i>	

Semi-Supervised and Transductive Learning

Bidirectional Semi-supervised Learning with Graphs	293
<i>Tomoharu Iwata and Kevin Duh</i>	
Coupled Bayesian Sets Algorithm for Semi-supervised Learning and Information Extraction	307
<i>Saurabh Verma and Estevam R. Hruschka Jr.</i>	
Graph-Based Transduction with Confidence	323
<i>Matan Orbach and Koby Crammer</i>	
Maximum Consistency Preferential Random Walks	339
<i>Deguang Kong and Chris Ding</i>	

Semi-supervised Multi-label Classification: A Simultaneous Large-Margin, Subspace Learning Approach	355
<i>Yuhong Guo and Dale Schuurmans</i>	

Sensor Data

MDL-Based Analysis of Time Series at Multiple Time-Scales	371
<i>Ugo Vespier, Arno Knobbe, Siegfried Nijssen, and Joaquin Vanschoren</i>	
Separable Approximate Optimization of Support Vector Machines for Distributed Sensing	387
<i>Sangkyun Lee, Marco Stolpe, and Katharina Morik</i>	
Unsupervised Inference of Auditory Attention from Biosensors	403
<i>Melih Kandemir, Arto Klami, Akos Vetek, and Samuel Kaski</i>	

Sequence and String Mining

A Family of Feed-Forward Models for Protein Sequence Classification . . .	419
<i>Sam Blasiak, Huzefa Rangwala, and Kathryn B. Laskey</i>	
General Algorithms for Mining Closed Flexible Patterns under Various Equivalence Relations	435
<i>Tomohiro I, Yuki Enokuma, Hideo Bannai, and Masayuki Takeda</i>	
Size Matters: Finding the Most Informative Set of Window Lengths . . .	451
<i>Jefrey Lijffijt, Panagiotis Papapetrou, and Kai Puolamäki</i>	

Social Network Mining

Discovering Links among Social Networks	467
<i>Francesco Buccafurri, Gianluca Lax, Antonino Nocera, and Domenico Ursino</i>	
Efficient Bi-objective Team Formation in Social Networks	483
<i>Mehdi Kargar, Aijun An, and Morteza Zihayat</i>	
Feature-Enhanced Probabilistic Models for Diffusion Network Inference	499
<i>Liaoruo Wang, Stefano Ermon, and John E. Hopcroft</i>	
Influence Spread in Large-Scale Social Networks – A Belief Propagation Approach	515
<i>Huy Nguyen and Rong Zheng</i>	

Location Affiliation Networks: Bonding Social and Spatial Information	531
<i>Konstantinos Pelechrinis and Prashant Krishnamurthy</i>	
On Approximation of Real-World Influence Spread	548
<i>Yu Yang, Enhong Chen, Qi Liu, Biao Xiang, Tong Xu, and Shafqat Ali Shad</i>	
Opinion Formation by Voter Model with Temporal Decay Dynamics	565
<i>Masahiro Kimura, Kazumi Saito, Kouzou Ohara, and Hiroshi Motoda</i>	
Viral Marketing for Product Cross-Sell through Social Networks	581
<i>Ramasuri Narayanam and Amit A. Nanavati</i>	
Which Topic Will You Follow?	597
<i>Deqing Yang, Yanghua Xiao, Bo Xu, Hanghang Tong, Wei Wang, and Sheng Huang</i>	

Spatial and Geographical Data Mining

Inferring Geographic Coincidence in Ephemeral Social Networks	613
<i>Honglei Zhuang, Alvin Chin, Sen Wu, Wei Wang, Xia Wang, and Jie Tang</i>	
Pedestrian Quantity Estimation with Trajectory Patterns	629
<i>Thomas Liebig, Zhao Xu, Michael May, and Stefan Wrobel</i>	
Socioscope: Spatio-temporal Signal Recovery from Social Media	644
<i>Jun-Ming Xu, Aniruddha Bhargava, Robert Nowak, and Xiaojin Zhu</i>	

Statistical Methods and Evaluation

A Framework for Evaluating the Smoothness of Data-Mining Results ...	660
<i>Gaurav Misra, Behzad Golshan, and Evimaria Terzi</i>	
Active Evaluation of Ranking Functions Based on Graded Relevance ...	676
<i>Christoph Sawade, Steffen Bickel, Timo von Oertzen, Tobias Scheffer, and Niels Landwehr</i>	

Time Series and Temporal Data Mining

Community Trend Outlier Detection Using Soft Temporal Pattern Mining	692
<i>Manish Gupta, Jing Gao, Yizhou Sun, and Jiawei Han</i>	
Data Structures for Detecting Rare Variations in Time Series	709
<i>Caio Valentim, Eduardo S. Laber, and David Sotelo</i>	

Invariant Time-Series Classification	725
<i>Josif Grabocka, Alexandros Nanopoulos, and Lars Schmidt-Thieme</i>	

Learning Bi-clustered Vector Autoregressive Models	741
<i>Tzu-Kuo Huang and Jeff Schneider</i>	

Transfer Learning

Discriminative Factor Alignment across Heterogeneous Feature Space . . .	757
<i>Fangwei Hu, Tianqi Chen, Nathan N. Liu, Qiang Yang, and Yong Yu</i>	

Learning to Perceive Two-Dimensional Displays Using Probabilistic Grammars	773
<i>Nan Li, William W. Cohen, and Kenneth R. Koedinger</i>	

Transfer Spectral Clustering	789
<i>Wenhao Jiang and Fu-lai Chung</i>	

System Demonstrations Track

An Aspect-Lexicon Creation and Evaluation Tool for Sentiment Analysis Researchers	804
<i>Mus'ab Husaini, Ahmet Koçyiğit, Dilek Tapucu, Berrin Yanikoglu, and Yücel Saygın</i>	

Association Rule Mining Following the Web Search Paradigm	808
<i>Radek Škrabal, Milan Šimůnek, Stanislav Vojříř, Andrej Hazucha, Tomáš Marek, David Chudán, and Tomáš Kliegr</i>	

ASV Monitor: Creating Comparability of Machine Learning Methods for Content Analysis	812
<i>Andreas Niekler, Patrick Jähnichen, and Gerhard Heyer</i>	

ClowdFlows: A Cloud Based Scientific Workflow Platform	816
<i>Janez Kranjc, Vid Podpečan, and Nada Lavrač</i>	

Extracting Trajectories through an Efficient and Unifying Spatio-temporal Pattern Mining System	820
<i>Phan Nhat Hai, Dino Ienco, Pascal Poncelet, and Maguelonne Teisseire</i>	

Knowledge Discovery through Symbolic Regression with HeuristicLab	824
<i>Gabriel Kronberger, Stefan Wagner, Michael Kommenda, Andreas Beham, Andreas Scheibenpflug, and Michael Affenzeller</i>	

<i>OutRules: A Framework for Outlier Descriptions in Multiple Context Spaces</i>	828
<i>Emmanuel Müller, Fabian Keller, Sebastian Blanc, and Klemens Böhm</i>	
Scientific Workflow Management with ADAMS	833
<i>Peter Reutemann and Joaquin Vanschoren</i>	
TopicExplorer: Exploring Document Collections with Topic Models	838
<i>Alexander Hinneburg, Rico Preiss, and René Schröder</i>	
VIKAMINE – Open-Source: Subgroup Discovery, Pattern Mining, and Analytics	842
<i>Martin Atzmueller and Florian Lemmerich</i>	
Nectar Track	
Learning Submodular Functions	846
<i>Maria-Florina Balcan and Nicholas J.A. Harvey</i>	
Matrix Factorization as Search	850
<i>Kristian Kersting, Christian Bauckhage, Christian Thureau, and Mirwaes Wahabzada</i>	
Metal Binding in Proteins: Machine Learning Complements X-Ray Absorption Spectroscopy	854
<i>Marco Lippi, Andrea Passerini, Marco Punta, and Paolo Frasconi</i>	
Modelling Input Varying Correlations between Multiple Responses	858
<i>Andrew Gordon Wilson and Zoubin Ghahramani</i>	
Author Index	863

Table of Contents – Part I

Invited Talks

Machine Learning for Robotics	1
<i>Pieter Abbeel</i>	
Declarative Modeling for Machine Learning and Data Mining	2
<i>Luc De Raedt</i>	
Machine Learning Methods for Music Discovery and Recommendation	4
<i>Douglas Eck</i>	
Solving Problems with Visual Analytics: Challenges and Applications . . .	5
<i>Daniel Keim</i>	
Analyzing Text and Social Network Data with Probabilistic Models . . .	7
<i>Padhraic Smyth</i>	

Association Rules and Frequent Patterns

Discovering Descriptive Tile Trees: By Mining Optimal Geometric Subtiles	9
<i>Nikolaj Tatti and Jilles Vreeken</i>	
Efficient Discovery of Association Rules and Frequent Itemsets through Sampling with Tight Performance Guarantees	25
<i>Matteo Riondato and Eli Upfal</i>	
Smoothing Categorical Data	42
<i>Arno Siebes and René Kersten</i>	

Bayesian Learning and Graphical Models

An Experimental Comparison of Hybrid Algorithms for Bayesian Network Structure Learning	58
<i>Maxime Gasse, Alex Aussem, and Haytham Elghazel</i>	
Bayesian Network Classifiers with Reduced Precision Parameters	74
<i>Sebastian Tschiatschek, Peter Reinprecht, Manfred Mücke, and Franz Pernkopf</i>	
Combining Subjective Probabilities and Data in Training Markov Logic Networks	90
<i>Tivadar Pápai, Shalini Ghosh, and Henry Kautz</i>	

Score-Based Bayesian Skill Learning 106
Shengbo Guo, Scott Sanner, Thore Graepel, and Wray Buntine

Classification

A Note on Extending Generalization Bounds for Binary Large-Margin
Classifiers to Multiple Classes 122
Ürün Dogan, Tobias Glasmachers, and Christian Igel

Extension of the Rocchio Classification Method to Multi-modal
Categorization of Documents in Social Media 130
Amin Mantrach and Jean-Michel Renders

Label-Noise Robust Logistic Regression and Its Applications 143
Jakramate Bootkrajang and Ata Kabán

Sentiment Classification with Supervised Sequence Embedding 159
Dmitriy Bespalov, Yanjun Qi, Bing Bai, and Ali Shokoufandeh

The Bitvector Machine: A Fast and Robust Machine Learning
Algorithm for Non-linear Problems 175
Stefan Edelkamp and Martin Stommel

**Dimensionality Reduction, Feature Selection and
Extraction**

Embedding Monte Carlo Search of Features in Tree-Based Ensemble
Methods 191
Francis Maes, Pierre Geurts, and Louis Wehenkel

Hypergraph Spectra for Semi-supervised Feature Selection 207
Zhihong Zhang, Edwin R. Hancock, and Xiao Bai

Learning Neighborhoods for Metric Learning 223
Jun Wang, Adam Woznica, and Alexandros Kalousis

Massively Parallel Feature Selection: An Approach Based on Variance
Preservation 237
Zheng Zhao, James Cox, David Duling, and Warren Sarle

PCA, Eigenvector Localization and Clustering for Side-Channel Attacks
on Cryptographic Hardware Devices 253
*Dimitrios Mavroeidis, Lejla Batina, Twan van Laarhoven, and
Elena Marchiori*

Distance-Based Methods and Kernels

Classifying Stem Cell Differentiation Images by Information Distance . . .	269
<i>Xianglilan Zhang, Hongnan Wang, Tony J. Collins, Zhigang Luo, and Ming Li</i>	
Distance Metric Learning Revisited	283
<i>Qiong Cao, Yiming Ying, and Peng Li</i>	
Geodesic Analysis on the Gaussian RKHS Hypersphere	299
<i>Nicolas Courty, Thomas Burger, and Pierre-François Marteau</i>	

Ensemble Methods

Boosting Nearest Neighbors for the Efficient Estimation of Posteriors . . .	314
<i>Roberto D'Ambrosio, Richard Nock, Wafa Bel Haj Ali, Frank Nielsen, and Michel Barlaud</i>	
Diversity Regularized Ensemble Pruning	330
<i>Nan Li, Yang Yu, and Zhi-Hua Zhou</i>	
Ensembles on Random Patches	346
<i>Gilles Louppe and Pierre Geurts</i>	

Graph and Tree Mining

An Efficiently Computable Support Measure for Frequent Subgraph Pattern Mining	362
<i>Yuyi Wang and Jan Ramon</i>	
Efficient Graph Kernels by Randomization	378
<i>Marion Neumann, Novi Patricia, Roman Garnett, and Kristian Kersting</i>	
Graph Mining for Object Tracking in Videos	394
<i>Fabien Diot, Elisa Fromont, Baptiste Jeudy, Emmanuel Marilly, and Olivier Martinot</i>	
Hypergraph Learning with Hyperedge Expansion	410
<i>Li Pu and Boi Faltings</i>	
Nearly Exact Mining of Frequent Trees in Large Networks	426
<i>Ashraf M. Kibriya and Jan Ramon</i>	
Reachability Analysis and Modeling of Dynamic Event Networks	442
<i>Kathy Macropol and Ambuj Singh</i>	

Large-Scale, Distributed and Parallel Mining and Learning

CC-MR – Finding Connected Components in Huge Graphs with MapReduce	458
<i>Thomas Seidl, Brigitte Boden, and Sergej Fries</i>	
Fast Near Neighbor Search in High-Dimensional Binary Data	474
<i>Anshumali Shrivastava and Ping Li</i>	
Fully Sparse Topic Models	490
<i>Khoat Than and Tu Bao Ho</i>	
Learning Compact Class Codes for Fast Inference in Large Multi Class Classification	506
<i>M. Cissé, T. Artières, and Patrick Gallinari</i>	
ParCube: Sparse Parallelizable Tensor Decompositions	521
<i>Evangelos E. Papalexakis, Christos Faloutsos, and Nicholas D. Sidiropoulos</i>	
Stochastic Coordinate Descent Methods for Regularized Smooth and Nonsmooth Losses	537
<i>Qing Tao, Kang Kong, Dejun Chu, and Gaowei Wu</i>	
Sublinear Algorithms for Penalized Logistic Regression in Massive Datasets	553
<i>Haoruo Peng, Zhengyu Wang, Edward Y. Chang, Shuchang Zhou, and Zhihua Zhang</i>	

Multi-Relational Mining and Learning

Author Name Disambiguation Using a New Categorical Distribution Similarity	569
<i>Shaohua Li, Gao Cong, and Chunyan Miao</i>	
Lifted Online Training of Relational Models with Stochastic Gradient Methods	585
<i>Babak Ahmadi, Kristian Kersting, and Sriraam Natarajan</i>	
Scalable Relation Prediction Exploiting Both Intrarelational Correlation and Contextual Information	601
<i>Xueyan Jiang, Volker Tresp, Yi Huang, Maximilian Nickel, and Hans-Peter Kriegel</i>	
Relational Differential Prediction	617
<i>Houssam Nassif, Vítor Santos Costa, Elizabeth S. Burnside, and David Page</i>	

Multi-Task Learning

Efficient Training of Graph-Regularized Multitask SVMs	633
<i>Christian Widmer, Marius Kloft, Nico Görnitz, and Gunnar Rätsch</i>	
Geometry Preserving Multi-task Metric Learning	648
<i>Peipei Yang, Kaizhu Huang, and Cheng-Lin Liu</i>	
Learning and Inference in Probabilistic Classifier Chains with Beam Search	665
<i>Abhishek Kumar, Shankar Vembu, Aditya Krishna Menon, and Charles Elkan</i>	
Learning Multiple Tasks with Boosted Decision Trees	681
<i>Jean Baptiste Faddoul, Boris Chidlovskii, Rémi Gilleron, and Fabien Torre</i>	
Multi-Task Boosting by Exploiting Task Relationships	697
<i>Yu Zhang and Dit-Yan Yeung</i>	
Sparse Gaussian Processes for Multi-task Learning	711
<i>Yuyang Wang and Roni Kharden</i>	

Natural Language Processing

Collective Information Extraction with Context-Specific Consistencies	728
<i>Peter Kluegl, Martin Toepfer, Florian Lemmerich, Andreas Hotho, and Frank Puppe</i>	
Supervised Learning of Semantic Relatedness	744
<i>Ran El-Yaniv and David Yanay</i>	
Unsupervised Bayesian Part of Speech Inference with Particle Gibbs	760
<i>Gregory Dubbin and Phil Blunsom</i>	
WikiSent: Weakly Supervised Sentiment Analysis through Extractive Summarization with Wikipedia	774
<i>Subhabrata Mukherjee and Pushpak Bhattacharyya</i>	

Online Learning and Data Streams

Adaptive Two-View Online Learning for Math Topic Classification	794
<i>Tam T. Nguyen, Kuiyu Chang, and Siu Cheung Hui</i>	
BDUOL: Double Updating Online Learning on a Fixed Budget	810
<i>Peilin Zhao and Steven C.H. Hoi</i>	

Handling Time Changing Data with Adaptive Very Fast Decision
Rules 827
 Petr Kosina and João Gama

Improved Counter Based Algorithms for Frequent Pairs Mining
in Transactional Data Streams 843
 Konstantin Kutzkov

Mirror Descent for Metric Learning: A Unified Approach 859
 Gautam Kunapuli and Jude Shavlik

Author Index 875