

Contents – Part I

Research Track

Classification, Regression and Supervised Learning

Data Split Strategies for Evolving Predictive Models	3
<i>Vikas C. Raykar and Amrita Saha</i>	
Discriminative Interpolation for Classification of Functional Data	20
<i>Rana Haber, Anand Rangarajan, and Adrian M. Peter</i>	
Fast Label Embeddings via Randomized Linear Algebra	37
<i>Paul Mineiro and Nikos Karampatziakis</i>	
Maximum Entropy Linear Manifold for Learning Discriminative Low-Dimensional Representation	52
<i>Wojciech Marian Czarnecki, Rafal Jozefowicz, and Jacek Tabor</i>	
Novel Decompositions of Proper Scoring Rules for Classification: Score Adjustment as Precursor to Calibration	68
<i>Meelis Kull and Peter Flach</i>	
Parameter Learning of Bayesian Network Classifiers Under Computational Constraints	86
<i>Sebastian Tschiatschek and Franz Pernkopf</i>	
Predicting Unseen Labels Using Label Hierarchies in Large-Scale Multi-label Learning	102
<i>Jinseok Nam, Eneldo Loza Mencía, Hyunwoo J. Kim, and Johannes Fürnkranz</i>	
Regression with Linear Factored Functions.	119
<i>Wendelin Böhmer and Klaus Obermayer</i>	
Ridge Regression, Hubness, and Zero-Shot Learning	135
<i>Yutaro Shigeto, Ikumi Suzuki, Kazuo Hara, Masashi Shimbo, and Yuji Matsumoto</i>	
Solving Prediction Games with Parallel Batch Gradient Descent	152
<i>Michael Großhans and Tobias Scheffer</i>	
Structured Regularizer for Neural Higher-Order Sequence Models.	168
<i>Martin Ratajczak, Sebastian Tschiatschek, and Franz Pernkopf</i>	

Versatile Decision Trees for Learning over Multiple Contexts.	184
<i>Reem Al-Otaibi, Ricardo B.C. Prudêncio, Meelis Kull, and Peter Flach</i>	
When Is Undersampling Effective in Unbalanced Classification Tasks?	200
<i>Andrea Dal Pozzolo, Olivier Caelen, and Gianluca Bontempi</i>	

Clustering and Unsupervised Learning

A Kernel-Learning Approach to Semi-supervised Clustering with Relative Distance Comparisons	219
<i>Ehsan Amid, Aristides Gionis, and Antti Ukkonen</i>	
Bayesian Active Clustering with Pairwise Constraints	235
<i>Yuanli Pei, Li-Ping Liu, and Xiaoli Z. Fern</i>	
ConDist: A Context-Driven Categorical Distance Measure	251
<i>Markus Ring, Florian Otto, Martin Becker, Thomas Niebler, Dieter Landes, and Andreas Hotho</i>	
Discovering Opinion Spammer Groups by Network Footprints	267
<i>Junting Ye and Leman Akoglu</i>	
Gamma Process Poisson Factorization for Joint Modeling of Network and Documents.	283
<i>Ayan Acharya, Dean Teffer, Jette Henderson, Marcus Tyler, Mingyuan Zhou, and Joydeep Ghosh</i>	
Generalization in Unsupervised Learning	300
<i>Karim T. Abou-Moustafa and Dale Schuurmans</i>	
Multiple Incomplete Views Clustering via Weighted Nonnegative Matrix Factorization with $L_{2,1}$ Regularization	318
<i>Weixiang Shao, Lifang He, and Philip S. Yu</i>	
Solving a Hard Cutting Stock Problem by Machine Learning and Optimisation.	335
<i>Steven D. Prestwich, Adejuyigbe O. Fajemisin, Laura Climent, and Barry O’Sullivan</i>	

Data Preprocessing

Markov Blanket Discovery in Positive-Unlabelled and Semi-supervised Data	351
<i>Konstantinos Sechidis and Gavin Brown</i>	
Multi-View Semantic Learning for Data Representation	367
<i>Peng Luo, Jinye Peng, Ziyu Guan, and Jianping Fan</i>	

Unsupervised Feature Analysis with Class Margin Optimization	383
<i>Sen Wang, Feiping Nie, Xiaojun Chang, Lina Yao, Xue Li, and Quan Z. Sheng</i>	

Data Streams and Online Learning

Ageing-Based Multinomial Naive Bayes Classifiers over Opinionated Data Streams	401
<i>Sebastian Wagner, Max Zimmermann, Eirini Ntoutsi, and Myra Spiliopoulou</i>	
Drift Detection Using Stream Volatility	417
<i>David Tse Jung Huang, Yun Sing Koh, Gillian Dobbie, and Albert Bifet</i>	
Early Classification of Time Series as a Non Myopic Sequential Decision Making Problem	433
<i>Asma Dachraoui, Alexis Bondu, and Antoine Cornuéjols</i>	
Ising Bandits with Side Information	448
<i>Shaona Ghosh and Adam Prügel-Bennett</i>	
Refined Algorithms for Infinitely Many-Armed Bandits with Deterministic Rewards.	464
<i>Yahel David and Nahum Shimkin</i>	

Deep Learning

An Empirical Investigation of Minimum Probability Flow Learning Under Different Connectivity Patterns	483
<i>Daniel Jiwoong Im, Ethan Buchman, and Graham W. Taylor</i>	
Difference Target Propagation	498
<i>Dong-Hyun Lee, Saizheng Zhang, Asja Fischer, and Yoshua Bengio</i>	
Online Learning of Deep Hybrid Architectures for Semi-supervised Categorization.	516
<i>Alexander G. Ororbia II, David Reitter, Jian Wu, and C. Lee Giles</i>	
Scoring and Classifying with Gated Auto-Encoders	533
<i>Daniel Jiwoong Im and Graham W. Taylor</i>	
Sign Constrained Rectifier Networks with Applications to Pattern Decompositions	546
<i>Senjian An, QiuHong Ke, Mohammed Bennamoun, Farid Boussaid, and Ferdous Sohel</i>	

Aggregation Under Bias: Rényi Divergence Aggregation and its Implementation via Machine Learning Markets	560
<i>Amos J. Storkey, Zhanxing Zhu, and Jinli Hu</i>	

Distance and Metric Learning

Higher Order Fused Regularization for Supervised Learning with Grouped Parameters	577
<i>Koh Takeuchi, Yoshinobu Kawahara, and Tomoharu Iwata</i>	
Joint Semi-supervised Similarity Learning for Linear Classification	594
<i>Maria-Irina Nicolae, Éric Gaussier, Amaury Habrard, and Marc Sebban</i>	
Learning Compact and Effective Distance Metrics with Diversity Regularization.	610
<i>Pengtao Xie</i>	
Scalable Metric Learning for Co-Embedding	625
<i>Farzaneh Mirzazadeh, Martha White, András György, and Dale Schuurmans</i>	

Large Scale Learning and Big Data

Adaptive Stochastic Primal-Dual Coordinate Descent for Separable Saddle Point Problems	645
<i>Zhanxing Zhu and Amos J. Storkey</i>	
Hash Function Learning via Codewords.	659
<i>Yinjie Huang, Michael Georgiopoulos, and Georgios C. Anagnostopoulos</i>	
HierCost: Improving Large Scale Hierarchical Classification with Cost Sensitive Learning.	675
<i>Anveshi Charuvaka and Huzefa Rangwala</i>	
Large Scale Optimization with Proximal Stochastic Newton-Type Gradient Descent	691
<i>Ziqiang Shi and Rujie Liu</i>	
Erratum to: Bayesian Active Clustering with Pairwise Constraints	E1
<i>Yuanli Pei, Li-Ping Liu, and Xiaoli Z. Fern</i>	
Erratum to: Scalable Metric Learning for Co-Embedding	E3
<i>Farzaneh Mirzazadeh, Martha White, András György, and Dale Schuurmans</i>	

Author Index	705
-------------------------------	-----

Contents – Part II

Research Track

Matrix and Tensor Analysis

BoostMF: Boosted Matrix Factorisation for Collaborative Ranking	3
<i>Nipa Chowdhury, Xiongcai Cai, and Cheng Luo</i>	
Convex Factorization Machines	19
<i>Mathieu Blondel, Akinori Fujino, and Naonori Ueda</i>	
Generalized Matrix Factorizations as a Unifying Framework for Pattern Set Mining: Complexity Beyond Blocks	36
<i>Pauli Miettinen</i>	
Scalable Bayesian Non-Negative Tensor Factorization for Massive Count Data.	53
<i>Changwei Hu, Piyush Rai, Changyou Chen, Matthew Harding, and Lawrence Carin</i>	
A Practical Approach to Reduce the Learning Bias Under Covariate Shift . . .	71
<i>Van-Tinh Tran and Alex Aussem</i>	
Hyperparameter Optimization with Factorized Multilayer Perceptrons	87
<i>Nicolas Schilling, Martin Wistuba, Lucas Drumond, and Lars Schmidt-Thieme</i>	
Hyperparameter Search Space Pruning – A New Component for Sequential Model-Based Hyperparameter Optimization	104
<i>Martin Wistuba, Nicolas Schilling, and Lars Schmidt-Thieme</i>	
Multi-Task Learning with Group-Specific Feature Space Sharing	120
<i>Niloofar Yousefi, Michael Georgiopoulos, and Georgios C. Anagnostopoulos</i>	
Opening the Black Box: Revealing Interpretable Sequence Motifs in Kernel-Based Learning Algorithms	137
<i>Marina M.-C. Vidovic, Nico Görnitz, Klaus-Robert Müller, Gunnar Rätsch, and Marius Kloft</i>	

Pattern and Sequence Mining

Fast Generation of Best Interval Patterns for Nonmonotonic Constraints	157
<i>Aleksey Buzmakov, Sergei O. Kuznetsov, and Amedeo Napoli</i>	

Non-Parametric Jensen-Shannon Divergence.	173
<i>Hoang-Vu Nguyen and Jilles Vreeken</i>	
Swap Randomization of Bases of Sequences for Mining Satellite Image Times Series.	190
<i>Nicolas M��ger, Christophe Rigotti, and Catherine Pothier</i>	
The Difference and the Norm Characterising Similarities and Differences Between Databases	206
<i>Kailash Budhathoki and Jilles Vreeken</i>	
Preference Learning and Label Ranking	
Dyad Ranking Using A Bilinear Plackett-Luce Model	227
<i>Dirk Sch��fer and Eyke H��llermeier</i>	
Fast Training of Support Vector Machines for Survival Analysis.	243
<i>Sebastian P��lsterl, Nassir Navab, and Amin Katouzian</i>	
Superset Learning Based on Generalized Loss Minimization.	260
<i>Eyke H��llermeier and Weiwei Cheng</i>	
Probabilistic, Statistical, and Graphical Approaches	
Bayesian Modelling of the Temporal Aspects of Smart Home Activity with Circular Statistics.	279
<i>Tom Diethe, Niall Twomey, and Peter Flach</i>	
Message Scheduling Methods for Belief Propagation.	295
<i>Christian Knoll, Michael Rath, Sebastian Tschatschek, and Franz Pernkopf</i>	
Output-Sensitive Adaptive Metropolis-Hastings for Probabilistic Programs . . .	311
<i>David Tolpin, Jan-Willem van de Meent, Brooks Paige, and Frank Wood</i>	
Planning in Discrete and Continuous Markov Decision Processes by Probabilistic Programming.	327
<i>Davide Nitti, Vaishak Belle, and Luc De Raedt</i>	
Simplifying, Regularizing and Strengthening Sum-Product Network Structure Learning.	343
<i>Antonio Vergari, Nicola Di Mauro, and Floriana Esposito</i>	
Sparse Bayesian Recurrent Neural Networks.	359
<i>Sotirios P. Chatzis</i>	

Structured Prediction of Sequences and Trees Using Infinite Contexts	373
<i>Ehsan Shareghi, Gholamreza Haffari, Trevor Cohn, and Ann Nicholson</i>	
Temporally Coherent Role-Topic Models (TCRTM): Deinterlacing Overlapping Activity Patterns	390
<i>Evgeniy Bart, Bob Price, and John Hanley</i>	
The Blind Leading the Blind: Network-Based Location Estimation Under Uncertainty	406
<i>Eric Malmi, Arno Solin, and Aristides Gionis</i>	
Weighted Rank Correlation: A Flexible Approach Based on Fuzzy Order Relations.	422
<i>Sascha Henzgen and Eyke Hüllermeier</i>	

Rich Data

Concurrent Inference of Topic Models and Distributed Vector Representations.	441
<i>Debakar Shamanta, Sheikh Motahar Naim, Parang Saraf, Naren Ramakrishnan, and M. Shahriar Hossain</i>	
Differentially Private Analysis of Outliers	458
<i>Rina Okada, Kazuto Fukuchi, and Jun Sakuma</i>	
Inferring Unusual Crowd Events from Mobile Phone Call Detail Records . . .	474
<i>Yuxiao Dong, Fabio Pinelli, Yiannis Gkoufas, Zubair Nabi, Francesco Calabrese, and Nitesh V. Chawla</i>	
Learning Pretopological Spaces for Lexical Taxonomy Acquisition	493
<i>Guillaume Cleuziou and Gaël Dias</i>	
Multidimensional Prediction Models When the Resolution Context Changes.	509
<i>Adolfo Martínez-Usó and José Hernández-Orallo</i>	
Semi-Supervised Subspace Co-Projection for Multi-class Heterogeneous Domain Adaptation	525
<i>Min Xiao and Yuhong Guo</i>	
Towards Computation of Novel Ideas from Corpora of Scientific Text	541
<i>Haixia Liu, James Goulding, and Tim Brailsford</i>	

Social and Graphs

Discovering Audience Groups and Group-Specific Influencers	559
<i>Shuyang Lin, Qingbo Hu, Jingyuan Zhang, and Philip S. Yu</i>	

Estimating Potential Customers Anywhere and Anytime Based on Location-Based Social Networks	576
<i>Hsun-Ping Hsieh, Cheng-Te Li, and Shou-De Lin</i>	
Exact Hybrid Covariance Thresholding for Joint Graphical Lasso	593
<i>Qingming Tang, Chao Yang, Jian Peng, and Jinbo Xu</i>	
Fast Inbound Top-K Query for Random Walk with Restart	608
<i>Chao Zhang, Shan Jiang, Yucheng Chen, Yidan Sun, and Jiawei Han</i>	
Finding Community Topics and Membership in Graphs	625
<i>Matt Revelle, Carlotta Domeniconi, Mack Sweeney, and Aditya Johri</i>	
Finding Dense Subgraphs in Relational Graphs	641
<i>Vinay Jethava and Niko Beerenwinkel</i>	
Generalized Modularity for Community Detection	655
<i>Mohadeseh Ganji, Abbas Seifi, Hosein Alizadeh, James Bailey, and Peter J. Stuckey</i>	
Handling Oversampling in Dynamic Networks Using Link Prediction	671
<i>Benjamin Fish and Rajmonda S. Caceres</i>	
Hierarchical Sparse Dictionary Learning	687
<i>Xiao Bian, Xia Ning, and Geoff Jiang</i>	
Latent Factors Meet Homophily in Diffusion Modelling	701
<i>Minh-Duc Luu and Ee-Peng Lim</i>	
Maintaining Sliding-Window Neighborhood Profiles in Interaction Networks	719
<i>Rohit Kumar, Toon Calders, Aristides Gionis, and Nikolaj Tatti</i>	
Response-Guided Community Detection: Application to Climate Index Discovery	736
<i>Gonzalo A. Bello, Michael Angus, Navya Pedemane, Jitendra K. Harlalka, Fredrick H.M. Semazzi, Vipin Kumar, and Nagiza F. Samatova</i>	
Robust Classification of Information Networks by Consistent Graph Learning	752
<i>Shi Zhi, Jiawei Han, and Quanquan Gu</i>	
Author Index	769

Contents – Part III

Industrial Track

Autonomous HVAC Control, a Reinforcement Learning Approach	3
<i>Enda Barrett and Stephen Linder</i>	
Clustering by Intent: A Semi-Supervised Method to Discover Relevant Clusters Incrementally	20
<i>George Forman, Hila Nachlieli, and Renato Keshet</i>	
Country-Scale Exploratory Analysis of Call Detail Records Through the Lens of Data Grid Models	37
<i>Romain Guigourès, Dominique Gay, Marc Boullé, Fabrice Clérot, and Fabrice Rossi</i>	
Early Detection of Fraud Storms in the Cloud	53
<i>Hani Neuirth, Yehuda Finkelstein, Amit Hilbuch, Shai Nahum, Daniel Alon, and Elad Yom-Tov</i>	
Flexible Sliding Windows for Kernel Regression Based Bus Arrival Time Prediction	68
<i>Hoang Thanh Lam and Eric Bouillet</i>	
Learning Detector of Malicious Network Traffic from Weak Labels	85
<i>Vojtech Franc, Michal Sofka, and Karel Bartos</i>	
Online Analysis of High-Volume Data Streams in Astroparticle Physics	100
<i>Christian Bockermann, Kai Brügge, Jens Buss, Alexey Egorov, Katharina Morik, Wolfgang Rhode, and Tim Ruhe</i>	
Robust Representation for Domain Adaptation in Network Security.	116
<i>Karel Bartos and Michal Sofka</i>	
Safe Exploration for Active Learning with Gaussian Processes	133
<i>Jens Schreiter, Duy Nguyen-Tuong, Mona Eberts, Bastian Bischoff, Heiner Markert, and Marc Toussaint</i>	
Semi-Supervised Consensus Clustering for ECG Pathology Classification . . .	150
<i>Helena Aidos, André Lourenço, Diana Batista, Samuel Rota Bulò, and Ana Fred</i>	
Two Step Graph-Based Semi-supervised Learning for Online Auction Fraud Detection	165
<i>Phiradet Bangcharoensap, Hayato Kobayashi, Nobuyuki Shimizu, Satoshi Yamauchi, and Tsuyoshi Murata</i>	

Watch-It-Next: A Contextual TV Recommendation System	180
<i>Michal Aharon, Eshcar Hillel, Amit Kagian, Ronny Lempel, Hayim Makabee, and Raz Nissim</i>	

Nectar Track

Bayesian Hypothesis Testing in Machine Learning	199
<i>Giorgio Corani, Alessio Benavoli, Francesca Mangili, and Marco Zaffalon</i>	
Data-Driven Exploration of Real-Time Geospatial Text Streams	203
<i>Harald Bosch, Robert Krüger, and Dennis Thom</i>	
Discovering Neutrinos Through Data Analytics.	208
<i>Mathis Börner, Wolfgang Rhode, Tim Ruhe, and Katharina Morik</i>	
Listener-Aware Music Recommendation from Sensor and Social Media Data	213
<i>Markus Schedl</i>	
Logic-Based Incremental Process Mining	218
<i>Stefano Ferilli, Domenico Redavid, and Floriana Esposito</i>	
Mobility Mining for Journey Planning in Rome	222
<i>Michele Berlingerio, Veli Bicer, Adi Botea, Stefano Braghin, Nuno Lopes, Riccardo Guidotti, and Francesca Pratesi</i>	
Pattern Structures and Concept Lattices for Data Mining and Knowledge Processing.	227
<i>Mehdi Kaytoue, Victor Codocedo, Aleksey Buzmakov, Jaume Baixeries, Sergei O. Kuznetsov, and Amedeo Napoli</i>	
Privacy Preserving Blocking and Meta-Blocking	232
<i>Alexandros Karakasidis, Georgia Koloniari, and Vassilios S. Verykios</i>	
Social Data Mining and Seasonal Influenza Forecasts: The FluOutlook Platform	237
<i>Qian Zhang, Corrado Gioannini, Daniela Paolotti, Nicola Perra, Daniela Perrotta, Marco Quaggiotto, Michele Tizzoni, and Alessandro Vespignani</i>	
Star Classification Under Data Variability: An Emerging Challenge in Astrominformatics	241
<i>Ricardo Vilalta, Kinjal Dhar Gupta, and Ashish Mahabal</i>	
The Evolution of Social Relationships and Strategies Across the Lifespan . . .	245
<i>Yuxiao Dong, Nitesh V. Chawla, Jie Tang, Yang Yang, and Yang Yang</i>	

Understanding Where Your Classifier Does (Not) Work	250
<i>Wouter Duivesteijn and Julia Thaele</i>	
Visual Analytics Methodology for Scalable and Privacy-Respectful Discovery of Place Semantics from Episodic Mobility Data	254
<i>Natalia Andrienko, Gennady Andrienko, Georg Fuchs, and Piotr Jankowski</i>	
Will This Paper Increase Your <i>h</i> -index?	259
<i>Yuxiao Dong, Reid A. Johnson, and Nitesh V. Chawla</i>	
Demo Track	
CardioWheel: ECG Biometrics on the Steering Wheel	267
<i>André Lourenço, Ana Priscila Alves, Carlos Carreiras, Rui Policarpo Duarte, and Ana Fred</i>	
Data Patterns Explained with Linked Data	271
<i>Ilaria Tididi, Mathieu d'Aquin, and Enrico Motta</i>	
GAZOUILLE: Detecting and Illustrating Local Events from Geolocalized Social Media Streams	276
<i>Pierre Houdyer, Albrecht Zimmerman, Mehdi Kaytoue, Marc Plantevit, Joseph Mitchell, and Céline Robardet</i>	
HiDER: Query-Driven Entity Resolution for Historical Data.	281
<i>Bijan Ranjbar-Sahraei, Julia Efremova, Hossein Rahmani, Toon Calders, Karl Tuyls, and Gerhard Weiss</i>	
IICE: Web Tool for Automatic Identification of Chemical Entities and Interactions	285
<i>Andre Lamurias, Luka A. Clarke, and Francisco M. Couto</i>	
Interactively Exploring Supply and Demand in the UK Independent Music Scene.	289
<i>Matt McVicar, Cédric Mesnage, Jefrey Lijffijt, and Tijl De Bie</i>	
Monitoring Production Equipment in Remote Locations.	293
<i>Ujval Kamath, Ana Costa e Silva, and Michael O'Connell</i>	
MULTI-NETS: Web-Based Multilayer Network Visualization	298
<i>Matija Piškorec, Borut Sluban, and Tomislav Šmuc</i>	
OMEGA: An Order-Preserving SubMatrix Mining, Indexing and Search Tool	303
<i>Tao Jiang, Zhanhuai Li, Qun Chen, Zhong Wang, Kaiwen Li, and Wei Pan</i>	

Probabilistic Programming in Anglican	308
<i>David Tolpin, Jan-Willem van de Meent, and Frank Wood</i>	
ProbLog2: Probabilistic Logic Programming.	312
<i>Anton Dries, Angelika Kimmig, Wannes Meert, Joris Renkens, Guy Van den Broeck, Jonas Vlasselaer, and Luc De Raedt</i>	
Real Time Detection and Tracking of Spatial Event Clusters	316
<i>Natalia Andrienko, Gennady Andrienko, Georg Fuchs, Salvatore Rinzivillo, and Hans-Dieter Betz</i>	
S&P360: Multidimensional Perspective on Companies from Online Data Sources	320
<i>Michele Berlingiero, Stefano Braghin, Francesco Calabrese, Cody Dunne, Yiannis Gkoufas, Mauro Martino, Jamie Rasmussen, and Steven Ross</i>	
SCAVENGER – A Framework for Efficient Evaluation of Dynamic and Modular Algorithms	325
<i>Andrey Tyukin, Stefan Kramer, and Jörg Wicker</i>	
UrbanHubble: Location Prediction and Geo-Social Analytics in LBSN	329
<i>Roland Assam, Simon Feiden, and Thomas Seidl</i>	
VIPER – Visual Pattern Explorer	333
<i>Matthijs van Leeuwen and Lara Cardinaels</i>	
Visualization Support to Interactive Cluster Analysis	338
<i>Gennady Andrienko and Natalia Andrienko</i>	
Author Index	343

Machine Learning and Knowledge Discovery in
Databases

European Conference, ECML PKDD 2015, Porto,
Portugal, September 7-11, 2015, Proceedings, Part I
Appice, A.; Rodrigues, P.P.; Santos Costa, V.; Soares,
C.; Gama, J.; Jorge, A. (Eds.)

2015, LVIII, 709 p. 160 illus., Softcover

ISBN: 978-3-319-23527-1