

Lecture Notes in Artificial Intelligence

11051

Subseries of Lecture Notes in Computer Science

LNAI Series Editors

Randy Goebel

University of Alberta, Edmonton, Canada

Yuzuru Tanaka

Hokkaido University, Sapporo, Japan

Wolfgang Wahlster

DFKI and Saarland University, Saarbrücken, Germany

LNAI Founding Series Editor

Joerg Siekmann

DFKI and Saarland University, Saarbrücken, Germany

More information about this series at <http://www.springer.com/series/1244>

Michele Berlingerio · Francesco Bonchi
Thomas Gärtner · Neil Hurley
Georgiana Ifrim (Eds.)

Machine Learning and Knowledge Discovery in Databases

European Conference, ECML PKDD 2018
Dublin, Ireland, September 10–14, 2018
Proceedings, Part I

Editors

Michele Berlingerio
IBM Research - Ireland
Dublin, Ireland

Francesco Bonchi
Institute for Scientific Interchange
Turin, Italy

Thomas Gärtner
University of Nottingham
Nottingham, UK

Neil Hurley 
University College Dublin
Dublin, Ireland

Georgiana Ifrim 
University College Dublin
Dublin, Ireland

ISSN 0302-9743 ISSN 1611-3349 (electronic)
Lecture Notes in Artificial Intelligence
ISBN 978-3-030-10924-0 ISBN 978-3-030-10925-7 (eBook)
<https://doi.org/10.1007/978-3-030-10925-7>

Library of Congress Control Number: 2018965756

LNCS Sublibrary: SL7 – Artificial Intelligence

© Springer Nature Switzerland AG 2019

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

We are delighted to introduce the proceedings of the 2018 edition of the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML-PKDD 2018). The conference was held in Dublin, Ireland, during September 10–14, 2018. ECML-PKDD is an annual conference that provides an international forum for the discussion of the latest high-quality research results in all areas related to machine learning and knowledge discovery in databases, including innovative applications. This event is the premier European machine learning and data mining conference and builds upon a very successful series of ECML-PKDD conferences.

The scientific program was of high quality and consisted of technical presentations of accepted papers, plenary talks by distinguished keynote speakers, workshops, and tutorials. Accepted papers were organized in five different tracks:

- The Conference Track, which featured research contributions, presented as part of the main conference program
- The Journal Track, which featured papers that were reviewed and published separately in special issues of the Springer journals *Machine Learning* and *Data Mining and Knowledge Discovery*, and that were selected as suitable for presentation at the conference
- The Applied Data Science Track, which focused on the application of data science to practical real-world scenarios, including contributions from academia, industry, and non-governmental organizations
- The Demo Track, which presented working demonstrations of prototypes or fully operational systems that exploit data science techniques
- The Nectar Track, which presented an overview of recent scientific advances at the frontier of machine learning and data mining in conjunction with other disciplines, as published in related conferences and journals

In addition to these tracks, the conference also included a PhD Forum in which PhD students received constructive feedback on their research progress and interacted with their peers. Co-located with the conference, this year there were 17 workshops on related research topics and six tutorial presentations.

In total, 95% of the accepted papers in the conference have accompanying software and/or data and are flagged as Reproducible Research papers in the proceedings. This speaks to the growing importance of reproducible research for the ECML-PKDD community. In the online proceedings each Reproducible Research paper has a link to the code and data made available with the paper. We believe this is a tremendous resource for the community and hope to see this trend maintained over the coming years.

We are very happy with the continued interest from the research community in our conference. We received 353 papers for the main conference track, of which 94 were

accepted, yielding an acceptance rate of about 26%. This allowed us to define a very rich program with 94 presentations in the main conference track. Moreover, there was a 26% acceptance rate (143 submissions, 37 accepted) to the Applied Data Science Track and 15% to the Journal Track (151 submissions, 23 accepted). Including the Nectar Track papers and some Journal Track papers from the 2017 special issues that were held over for presentation until 2018, we had in total 166 parallel scientific talks during the three main conference days.

The program also included five plenary keynotes by invited speakers: Misha Bilenko (Head of Machine Intelligence and Research Yandex, Moscow, Russia), Corinna Cortes (Head of Google Research New York, USA), Aristides Gionis (Professor, Department of Computer Science, Aalto University, Helsinki, Finland), Cynthia Rudin (Associate Professor of Computer Science and Electrical and Computer Engineering, Duke University, Durham, North Carolina, USA), and Naftali Tishby (Professor, School of Engineering and Computer Science, Hebrew University of Jerusalem, Israel).

This year, ECML-PKDD attracted over 630 participants from 42 countries. It attracted substantial attention from industry both through sponsorship and submission/participation at the conference and workshops. Moreover, ECML-PKDD hosted a very popular Nokia Women in Science Luncheon to discuss the importance of awareness of equal opportunity and support for women in science and technology.

The Awards Committee selected research papers considered to be of exceptional quality and worthy of special recognition:

- ML Student Best Paper Award: “Hyperparameter Learning for Conditional Mean Embeddings with Rademacher Complexity Bounds,” by Kelvin Hsu, Richard Nock and Fabio Ramos
- KDD Student Best Paper Award: “Anytime Subgroup Discovery in Numerical Domains with Guarantees,” by Aimene Belfodil, Adnene Belfodil and Mehdi Kaytoue.

We would like to thank all participants, authors, reviewers, and organizers of the conference for their contribution to making ECML-PKDD 2018 a great scientific event.

We would also like to thank the Croke Park Conference Centre and the student volunteers. Thanks to Springer for their continuous support and Microsoft for allowing us to use their CMT software for conference management and providing support throughout. Special thanks to our many sponsors and to the ECML-PKDD Steering Committee for their support and advice. Finally, we would like to thank the organizing institutions: the Insight Centre for Data Analytics, University College Dublin, Ireland, and IBM Research, Ireland.

September 2018

Michele Berlingerio
 Francesco Bonchi
 Thomas Gärtner
 Georgiana Ifrim
 Neil Hurley

Organization

ECML PKDD 2018 Organization

General Chairs

Michele Berlingerio IBM Research, Ireland
Neil Hurley University College Dublin, Ireland

Program Chairs

Michele Berlingerio IBM Research, Ireland
Francesco Bonchi ISI Foundation, Italy
Thomas Gärtner University of Nottingham, UK
Georgiana Ifrim University College Dublin, Ireland

Journal Track Chairs

Björn Bringmann McKinsey & Company, Germany
Jesse Davis Katholieke Universiteit Leuven, Belgium
Elisa Fromont IRISA, Rennes 1 University, France
Derek Greene University College Dublin, Ireland

Applied Data Science Track Chairs

Edward Curry National University of Ireland Galway, Ireland
Alice Marascu Nokia Bell Labs, Ireland

Workshop and Tutorial Chairs

Carlos Alzate IBM Research, Ireland
Anna Monreale University of Pisa, Italy

Nectar Track Chairs

Ulf Brefeld Leuphana University of Lüneburg, Germany
Fabio Pinelli Vodafone, Italy

Demo Track Chairs

Elizabeth Daly IBM Research, Ireland
Brian Mac Namee University College Dublin, Ireland

PhD Forum Chairs

Bart Goethals University of Antwerp, Belgium
Dafna Shahaf Hebrew University of Jerusalem, Israel

Discovery Challenge Chairs

Martin Atzmüller Tilburg University, The Netherlands
Francesco Calabrese Vodafone, Italy

Awards Committee

Tijl De Bie Ghent University, Belgium
Arno Siebes Utrecht University, The Netherlands
Bart Goethals University of Antwerp, Belgium
Walter Daelemans University of Antwerp, Belgium
Katharina Morik TU Dortmund, Germany

Professional Conference Organizer

Keynote PCO Dublin, Ireland

www.keynotepco.ie

ECML PKDD Steering Committee

Michele Sebag Université Paris Sud, France
Francesco Bonchi ISI Foundation, Italy
Albert Bifet Télécom ParisTech, France
Hendrik Blockeel KU Leuven, Belgium and Leiden University, The Netherlands
Katharina Morik University of Dortmund, Germany
Arno Siebes Utrecht University, The Netherlands
Siegfried Nijssen LIACS, Leiden University, The Netherlands
Chedy Raïssi Inria Nancy Grand-Est, France
João Gama FCUP, University of Porto/LIAAD, INESC Porto L.A.,
Portugal
Annalisa Appice University of Bari Aldo Moro, Italy
Indré Žliobaitė University of Helsinki, Finland
Andrea Passerini University of Trento, Italy
Paolo Frasconi University of Florence, Italy
Céline Robardet National Institute of Applied Science Lyon, France

Jilles Vreeken	Saarland University, Max Planck Institute for Informatics, Germany
Sašo Džeroski	Jožef Stefan Institute, Slovenia
Michelangelo Ceci	University of Bari Aldo Moro, Italy
Myra Spiliopoulou	Magdeburg University, Germany
Jaakko Hollmén	Aalto University, Finland

Area Chairs

Michael Berthold	Universität Konstanz, Germany
Hendrik Blockeel	KU Leuven, Belgium and Leiden University, The Netherlands
Ulf Brefeld	Leuphana University of Lüneburg, Germany
Toon Calders	University of Antwerp, Belgium
Michelangelo Ceci	University of Bari Aldo Moro, Italy
Bruno Cremilleux	Université de Caen Normandie, France
Tapio Elomaa	Tampere University of Technology, Finland
Johannes Fürnkranz	TU Darmstadt, Germany
Peter Flach	University of Bristol, UK
Paolo Frasconi	University of Florence, Italy
João Gama	FCUP, University of Porto/LIAAD, INESC Porto L.A., Portugal
Jaakko Hollmén	Aalto University, Finland
Alípio Jorge	FCUP, University of Porto/LIAAD, INESC Porto L.A., Portugal
Stefan Kramer	Johannes Gutenberg University Mainz, Germany
Giuseppe Manco	ICAR-CNR, Italy
Siegfried Nijssen	LIACS, Leiden University, The Netherlands
Andrea Passerini	University of Trento, Italy
Arno Siebes	Utrecht University, The Netherlands
Myra Spiliopoulou	Magdeburg University, Germany
Luis Torgo	Dalhousie University, Canada
Celine Vens	KU Leuven, Belgium
Jilles Vreeken	Saarland University, Max Planck Institute for Informatics, Germany

Conference Track Program Committee

Carlos Alzate	Roberto Bayardo	Indrajit Bhattacharya
Aijun An	Martin Becker	Marenglen Biba
Fabrizio Angiulli	Srikanta Bedathur	Silvio Bicchato
Annalisa Appice	Jessa Bekker	Mario Boley
Ira Assent	Vaishak Belle	Gianluca Bontempi
Martin Atzmueller	Andras Benczur	Henrik Bostrom
Antonio Bahamonde	Daniel Bengs	Tassadit Bouadi
Jose Balcazar	Petr Berka	Pavel Brazdil

Dariusz Brzezinski	Cesar Ferri	Latifur Khan
Rui Camacho	Răzvan Florian	Frank Klawonn
Longbing Cao	Eibe Frank	Jiri Klema
Francisco Casacuberta	Elisa Fromont	Tomas Kliegr
Peggy Cellier	Fabio Fumarola	Marius Kloft
Loic Cerf	Esther Galbrun	Dragi Kocev
Tania Cerquitelli	Patrick Gallinari	Levente Kocsis
Edward Chang	Dragan Gamberger	Yun Sing Koh
Keke Chen	Byron Gao	Alek Kolcz
Weiwei Cheng	Paolo Garza	Irena Koprinska
Silvia Chiusano	Konstantinos Georgatzis	Frederic Koriche
Arthur Choi	Pierre Geurts	Walter Kusters
Frans Coenen	Dorota Glowacka	Lars Kotthoff
Mário Cordeiro	Nico Goernitz	Danai Koutra
Roberto Corizzo	Elsa Gomes	Georg Krempf
Vitor Santos Costa	Mehmet Gönen	Tomas Krilavicius
Bertrand Cuissart	James Goulding	Yamuna Krishnamurthy
Boris Cule	Michael Granitzer	Matjaz Kukar
Tomaž Curk	Caglar Gulcehre	Meelis Kull
James Cussens	Francesco Gullo	Prashanth L. A.
Alfredo Cuzzocrea	Stephan Günnemann	Jorma Laaksonen
Claudia d'Amato	Tias Guns	Nicolas Lachiche
Maria Damiani	Sara Hajian	Leo Lahti
Tijl De Bie	Maria Halkidi	Helge Langseth
Martine De Cock	Jiawei Han	Thomas Lansdall-Welfare
Juan Jose del Coz	Mohammad Hasan	Christine Largeton
Anne Denton	Xiao He	Pedro Larranaga
Christian Desrosiers	Denis Helic	Silvio Lattanzi
Nicola Di Mauro	Daniel Hernandez-Lobato	Niklas Lavesson
Claudia Diamantini	Jose Hernandez-Orallo	Binh Le
Uwe Dick	Thanh Lam Hoang	Freddy Lecue
Tom Diethe	Frank Hoepfner	Florian Lemmerich
Ivica Dimitrovski	Arjen Hommersom	Jiuyong Li
Wei Ding	Tamas Horvath	Limin Li
Ying Ding	Andreas Hotho	Jefrey Lijffijt
Stephan Doerfel	Yuanhua Huang	Tony Lindgren
Carlotta Domeniconi	Eyke Hüllermeier	Corrado Loglisci
Frank Dondelinger	Dino Ienco	Peter Lucas
Madalina Drugan	Szymon Jaroszewicz	Brian Mac Namee
Wouter Duivesteijn	Giuseppe Jurman	Gjorgji Madjarov
Inês Dutra	Toshihiro Kamishima	Sebastian Mair
Dora Erdos	Michael Kamp	Donato Malerba
Fabio Fassetto	Bo Kang	Luca Martino
Ad Feelders	Andreas Karwath	Elio Masciari
Stefano Ferilli	George Karypis	Andres Masegosa
Carlos Ferreira	Mehdi Kaytoue	Florent Masseglia

Ernestina Menasalvas	Chotirat Ratanamahatana	Domenico Talia
Corrado Mencar	Jan Rauch	Letizia Tanca
Rosa Meo	Chiara Renso	Jovan Tanevski
Pauli Miettinen	Achim Rettinger	Nikolaj Tatti
Dunja Mladenic	Fabrizio Riguzzi	Maguelonne Teisseire
Karthika Mohan	Matteo Riondato	Georgios Theocharous
Anna Monreale	Celine Robardet	Ljupco Todorovski
Joao Moreira	Juan Rodriguez	Roberto Trasarti
Mohamed Nadif	Fabrice Rossi	Volker Tresp
Ndapa Nakashole	Celine Rouveirol	Isaac Triguero
Jinseok Nam	Stefan Rueping	Panayiotis Tsaparas
Mirco Nanni	Salvatore Ruggieri	Vincent S. Tseng
Amedeo Napoli	Yvan Saeys	Karl Tuyls
Nicolo Navarin	Alan Said	Niall Twomey
Benjamin Negrevergne	Lorenza Saitta	Nikolaos Tziortziotis
Benjamin Nguyen	Tomoya Sakai	Theodoros Tzouramanis
Xia Ning	Alessandra Sala	Antti Ukkonen
Kjetil Norvag	Ansaf Salleb-Aouissi	Toon Van Craenendonck
Eirini Ntoutsis	Claudio Sartori	Martijn Van Otterlo
Andreas Nurnberger	Pierre Schaus	Iraklis Varlamis
Barry O'Sullivan	Lars Schmidt-Thieme	Julien Velcin
Dino Oglic	Christoph Schommer	Shankar Vembu
Francesco Orsini	Matthias Schubert	Deepak Venugopal
Nikunj Oza	Konstantinos Sechidis	Vassilios S. Verykios
Pance Panov	Sohan Seth	Ricardo Vigarío
Apostolos Papadopoulos	Vinay Setty	Herna Viktor
Panayiotis Papapetrou	Junming Shao	Christel Vrain
Ioannis Partalas	Nikola Simidjievski	Willem Waegeman
Gabriella Pasi	Sameer Singh	Jianyong Wang
Dino Pedreschi	Andrzej Skowron	Joerg Wicker
Jaakko Peltonen	Dominik Slezak	Marco Wiering
Ruggero Pensa	Kevin Small	Martin Wistuba
Iker Perez	Gavin Smith	Philip Yu
Nico Piatkowski	Tomislav Smuc	Bianca Zadrozny
Andrea Pietracaprina	Yangqiu Song	Gerson Zaverucha
Gianvito Pio	Arnaud Soulet	Bernard Zenko
Susanna Pirttikangas	Wesllen Sousa	Junping Zhang
Marc Plantevit	Alessandro Sperduti	Min-Ling Zhang
Pascal Poncelet	Jerzy Stefanowski	Shichao Zhang
Miguel Prada	Giovanni Stilo	Ying Zhao
Philippe Preux	Gerd Stumme	Mingjun Zhong
Buyue Qian	Mahito Sugiyama	Albrecht Zimmermann
Chedy Raissi	Mika Sulkava	Marinka Zitnik
Jan Ramon	Einoshin Suzuki	Indr� Zliobait�
Huzefa Rangwala	Stephen Swift	
Zbigniew Ras	Andrea Tagarelli	

Applied Data Science Track Program Committee

Oznur Alkan	Sidath Handurukande	Nikunj Oza
Carlos Alzate	Souleiman Hasan	Ioannis Partalas
Nicola Barberi	Georges Hebrail	Milan Petkovic
Gianni Barlacchi	Thanh Lam	Fabio Pinelli
Enda Barrett	Neil Hurley	Yongrui Qin
Roberto Bayardo	Hongxia Jin	Rene Quiniou
Srikanta Bedathur	Anup Kalia	Ambrish Rawat
Daniel Bengs	Pinar Karagoz	Fergal Reid
Cuissart Bertrand	Mehdi Kaytoue	Achim Rettinger
Urvesh Bhowan	Alek Kolcz	Stefan Rueping
Tassadit Bouadi	Deguano Kong	Elizeu Santos-Neto
Thomas Brovelli	Lars Kotthoff	Manali Sharma
Teodora Sandra	Nick Koudas	Alkis Simitsis
Berkant Barla	Hardy Kremer	Kevin Small
Michelangelo Ceci	Helge Langseth	Alessandro Sperduti
Edward Chang	Freddy Lecue	Siqi Sun
Soumyadeep Chatterjee	Zhenhui Li	Pal Sundsoy
Abon Chaudhuri	Lin Liu	Ingo Thon
Javier Cuenca	Jiebo Luo	Marko Tkalcić
Mahashweta Das	Arun Maiya	Luis Torgo
Viktoriya Degeler	Silviu Maniu	Radu Tudoran
Wei Ding	Elio Masciari	Umair Ul
Yuxiao Dong	Luis Matias	Jan Van
Carlos Ferreira	Dimitrios Mavroeidis	Ranga Vatsavai
Andre Freitas	Charalampos	Fei Wang
Feng Gao	Mavroforakis	Xiang Wang
Dinesh Garg	James McDermott	Wang Wei
Guillermo Garrido	Daniil Mirylenka	Martin Wistuba
Rumi Ghosh	Elena Mocanu	Erik Wittern
Martin Gleize	Raul Moreno	Milena Yankova
Slawek Goryczka	Bogdan Nicolae	Daniela Zaharie
Riccardo Guidotti	Maria-Irina Nicolae	Chongsheng Zhang
Francesco Gullo	Xia Ning	Yanchang Zhao
Thomas Guyet	Sean O'Riain	Albrecht Zimmermann
Allan Hanbury	Adegboyega Ojo	

Nectar Track Program Committee

Annalisa Appice	Peter Flach	Ernestina Menasalvas
Martin Atzmüller	Johannes Fürnkranz	Mirco Musolesi
Hendrik Blockeel	Joao Gama	Franco Maria
Ulf Brefeld	Andreas Hotho	Maryam Tavakol
Tijl De	Kristian Kersting	Gabriele Tolomei
Kurt Driessens	Sebastian Mair	Salvatore Trani

Demo Track Program Committee

Gustavo Carneiro
Derek Greene
Mark Last
Vincent Lemaire

Brian Mac Namee
Susan McKeever
Joao Papa
Niladri Sett

Konstantinos Skiannis
Jerzy Stefanowski
Luis Teixeira
Grigorios Tsoumakas

Invited Talks Abstracts

Building Production Machine Learning Systems

Misha Bilenko

Head of Machine Intelligence and Research, Yandex, Moscow, Russia

Abstract. How does one build a production ML system that can effectively incorporate corrections, while avoiding the typical risks and engineering costs of online learning methods? An effective solution requires a combination of very latest and some well-dated algorithms. Parametric machine learning methods - such as neural networks, boosted trees, factorization methods and their ensembles - yield state-of-the-art results on ML benchmarks and competitions. Real-world deployments of ML systems, however, differ dramatically from those static settings. We discuss issues that differentiate production and academic ML systems, leading to the need for combining parametric models with their non-parametric brethren, i.e, modern variants of Nearest Neighbor algorithms. The combined approach is particularly suitable for systems where incorporating corrections must be accomplished rapidly, as illustrated by some lively real-life examples from a large-scale conversational assistant.

Bio: Misha Bilenko heads the Machine Intelligence and Research (MIR) division at Yandex, which integrates research and product development in core AI areas: machine learning, dialog systems, speech recognition and synthesis, machine translation and computer vision. Before Yandex, he led the Machine Learning Algorithms team at Microsoft, which shipped ML technologies in multiple products across all Microsoft divisions. He started his career in the Machine Learning Group in Microsoft Research after receiving his Ph.D. in Computer Science from the University of Texas at Austin and stints at Google and IBM Research.

Combating Misinformation and Building Trust in the Media

Corinna Cortes

Head of Google Research, New York, USA

Abstract. Trust in journalism, search engines and social media has been on a sharp decline over the last years. “Fake News” is a reality and internet users struggle to distinguish credible information from falsehood. In this talk we will discuss tools and methodology we as researchers at Google are building to restore online trust and combat misinformation.

Bio: Corinna Cortes is the Head of Google Research, NY, where she is working on a broad range of theoretical and applied large-scale machine learning problems. Corinna speaks about how she got into the tech world, and some work that is happening within the Research team at Google. Prior to Google, Corinna spent more than ten years at AT&T Labs – Research, formerly AT&T Bell Labs, where she held a distinguished research position. Corinna’s research work is well-known in particular for her contributions to the theoretical foundations of support vector machines (SVMs), for which she jointly with Vladimir Vapnik received the 2008 Paris Kanellakis Theory and Practice Award, and her work on data-mining in very large data sets for which she was awarded the AT&T Science and Technology Medal in the year 2000. Corinna received her MS degree in Physics from University of Copenhagen and joined AT&T Bell Labs as a researcher in 1989. She received her Ph.D. in computer science from the University of Rochester in 1993. Corinna is also a competitive runner and a mother of twins.

Combating Bias and Polarization in Online Media

Aristides Gionis

Department of Computer Science, Aalto University, Helsinki, Finland

Abstract. Online social media are a major venue of public discourse today, hosting the opinions of hundreds of millions of individuals. Social media are often credited for providing a technological means to break information barriers and promote diversity and democracy. In practice, however, the opposite effect is often observed: users tend to favor content that agrees with their existing world-view, get less exposure to conflicting viewpoints, and eventually create “echo chambers” and increased polarization. Arguably, without any kind of moderation, current social-media platforms gravitate towards a state in which net-citizens are constantly reinforcing their existing opinions. In this talk we present an ongoing line of work on analyzing and moderating online social discussions. We first consider the questions of detecting controversy using network structure and content. We then address the problem of designing algorithms to break filter bubbles, reduce polarization, and increase diversity. We discuss a number of different strategies such as user and content recommendation, as well as approaches based on information cascades.

Bio: Aristides Gionis is a professor in the department of Computer Science in Aalto University. He has been a visiting professor in the University of Rome (2016) and a senior research scientist in Yahoo! Research (2006–2012). He is currently serving as an action editor in the Data Management and Knowledge Discovery journal (DMKD), an associate editor in the ACM Transactions on Knowledge Discovery from Data (TKDD), and a managing editor in Internet Mathematics. He has contributed in several areas of data science, such as algorithmic data analysis, web mining, social-media analysis, data clustering, and privacy-preserving data mining. His current research is funded by the Academy of Finland (projects Nestor, Agra, AIDA) and the European Commission (project SoBigData).

New Algorithms for Interpretable Machine Learning

Cynthia Rudin

Associate Professor of Computer Science and Electrical
and Computer Engineering
Duke University, Durham, North Carolina, USA

Abstract. What if every black box machine learning model could be replaced with one that was equally accurate but also interpretable? If we could do this, we would identify flaws in our models and data that we could not see before. Perhaps we could prevent some of the poor decisions in criminal justice and medicine that are caused by problems with using black box models. We could also eliminate the need for “explanations” that are misleading and often wrong. I will present algorithms for (i) interpretable neural networks for computer vision, (ii) certifiably optimal decision lists, and (iii) certifiably optimal scoring systems (sparse linear models with integer coefficients). Models from these algorithms can often be used in place of black box models, while achieving the same accuracy.

Bio: Cynthia Rudin is an associate professor of computer science, electrical and computer engineering, and statistics at Duke University, and directs the Prediction Analysis Lab. Previously, Prof. Rudin held positions at MIT, Columbia, and NYU. She holds an undergraduate degree from the University at Buffalo, and a PhD in applied and computational mathematics from Princeton University. She is the recipient of the 2013 and 2016 INFORMS Innovative Applications in Analytics Awards, an NSF CAREER award, was named as one of the “Top 40 Under 40” by Poets and Quants in 2015, and was named by Businessinsider.com as one of the 12 most impressive professors at MIT in 2015. Work from her lab has won 10 best paper awards in the last 5 years. She is past chair of the INFORMS Data Mining Section, and is currently chair of the Statistical Learning and Data Science section of the American Statistical Association. She also serves on (or has served on) committees for DARPA, the National Institute of Justice, the National Academy of Sciences (for both statistics and criminology/law), and AAAI.

Information Theory of Deep Learning

Naftali Tishby

School of Engineering and Computer Science, Hebrew University of Jerusalem,
Israel

Abstract. I will present new results on the organization and interpretability of the many layers in deep learning, which stem from my information bottleneck theory of deep neural networks.

Bio: Dr. Naftali Tishby is a professor of Computer Science, and the incumbent of the Ruth and Stan Flinkman Chair for Brain Research at the Edmond and Lily Safra Center for Brain Science (ELSC) at the Hebrew University of Jerusalem. He is one of the leaders of machine learning research and computational neuroscience in Israel and his numerous ex-students serve at key academic and industrial research positions all over the world. Prof. Tishby was the founding chair of the new computer-engineering program, and a director of the Leibnitz research center in computer science, at the Hebrew university. Tishby received his PhD in theoretical physics from the Hebrew university in 1985 and was a research staff member at MIT and Bell Labs from 1985 and 1991. Prof. Tishby was also a visiting professor at Princeton NECI, University of Pennsylvania, UCSB, and IBM research. His current research is at the interface between computer science, statistical physics, and computational neuroscience. He pioneered various applications of statistical physics and information theory in computational learning theory. More recently, he has been working on the foundations of biological information processing and the connections between dynamics and information. He has introduced with his colleagues new theoretical frameworks for optimal adaptation and efficient information representation in biology, such as the Information Bottleneck method and the Minimum Information principle for neural coding.

Contents – Part I

Adversarial Learning

Image Anomaly Detection with Generative Adversarial Networks	3
<i>Lucas Deecke, Robert Vandermeulen, Lukas Ruff, Stephan Mandt, and Marius Kloft</i>	
Image-to-Markup Generation via Paired Adversarial Learning.	18
<i>Jin-Wen Wu, Fei Yin, Yan-Ming Zhang, Xu-Yao Zhang, and Cheng-Lin Liu</i>	
Toward an Understanding of Adversarial Examples in Clinical Trials	35
<i>Konstantinos Papangelou, Konstantinos Sechidis, James Weatherall, and Gavin Brown</i>	
ShapeShifter: Robust Physical Adversarial Attack on Faster R-CNN Object Detector.	52
<i>Shang-Tse Chen, Cory Cornelius, Jason Martin, and Duen Horng (Polo) Chau</i>	

Anomaly and Outlier Detection

GRIDWATCH: Sensor Placement and Anomaly Detection in the Electrical Grid . . .	71
<i>Bryan Hooi, Dhivya Eswaran, Hyun Ah Song, Amritanshu Pandey, Marko Jereminov, Larry Pileggi, and Christos Faloutsos</i>	
Incorporating Privileged Information to Unsupervised Anomaly Detection . . .	87
<i>Shubhanshu Shekhar and Leman Akoglu</i>	
L1-Depth Revisited: A Robust Angle-Based Outlier Factor in High-Dimensional Space	105
<i>Ninh Pham</i>	
Beyond Outlier Detection: LOOKOUT for Pictorial Explanation.	122
<i>Nikhil Gupta, Dhivya Eswaran, Neil Shah, Leman Akoglu, and Christos Faloutsos</i>	
CONOUT: Contextual Outlier Detection with Multiple Contexts: Application to Ad Fraud	139
<i>M. Y. Meghanath, Deepak Pai, and Leman Akoglu</i>	
Scalable and Interpretable One-Class SVMs with Deep Learning and Random Fourier Features	157
<i>Minh-Nghia Nguyen and Ngo Anh Vien</i>	

Group Anomaly Detection Using Deep Generative Models.	173
<i>Raghavendra Chalapathy, Edward Toth, and Sanjay Chawla</i>	

Applications

Detecting Autism by Analyzing a Simulated Social Interaction	193
<i>Hanna Drimalla, Niels Landwehr, Irina Baskow, Behnoush Behnia, Stefan Roepke, Isabel Dziobek, and Tobias Scheffer</i>	
A Discriminative Model for Identifying Readers and Assessing Text Comprehension from Eye Movements	209
<i>Silvia Makowski, Lena A. Jäger, Ahmed Abdelwahab, Niels Landwehr, and Tobias Scheffer</i>	
Face-Cap: Image Captioning Using Facial Expression Analysis.	226
<i>Omid Mohamad Nezami, Mark Dras, Peter Anderson, and Len Hamey</i>	
Pedestrian Trajectory Prediction with Structured Memory Hierarchies	241
<i>Tharindu Fernando, Simon Denman, Sridha Sridharan, and Clinton Fookes</i>	

Classification

Multiple Instance Learning with Bag-Level Randomized Trees	259
<i>Tomáš Komárek and Petr Somol</i>	
One-Class Quantification	273
<i>Denis dos Reis, André Maletzke, Everton Cherman, and Gustavo Batista</i>	
Deep F-Measure Maximization in Multi-label Classification: A Comparative Study	290
<i>Stijn Decubber, Thomas Mortier, Krzysztof Dembczyński, and Willem Waegeman</i>	
Ordinal Label Proportions	306
<i>Rafael Poyiadzi, Raúl Santos-Rodríguez, and Tijl De Bie</i>	
AWX: An Integrated Approach to Hierarchical-Multilabel Classification	322
<i>Luca Masera and Enrico Blanzieri</i>	

Clustering and Unsupervised Learning

Clustering in the Presence of Concept Drift	339
<i>Richard Hugh Moulton, Herna L. Viktor, Nathalie Japkowicz, and João Gama</i>	

Time Warp Invariant Dictionary Learning for Time Series Clustering:
 Application to Music Data Stream Analysis 356
*Saeed Varasteh Yazdi, Ahlame Douzal-Chouakria, Patrick Gallinari,
 and Manuel Moussallam*

How Your Supporters and Opponents Define Your Interestingness 373
Bruno Crémilleux, Arnaud Giacometti, and Arnaud Soulet

Deep Learning

Efficient Decentralized Deep Learning by Dynamic Model Averaging 393
*Michael Kamp, Linara Adilova, Joachim Sicking, Fabian Hüger,
 Peter Schlicht, Tim Wirtz, and Stefan Wrobel*

Using Supervised Pretraining to Improve Generalization of Neural
 Networks on Binary Classification Problems. 410
*Alex Yuxuan Peng, Yun Sing Koh, Patricia Riddle,
 and Bernhard Pfahringer*

Towards Efficient Forward Propagation on Resource-Constrained Systems . . . 426
*Günther Schindler, Matthias Zöhrer, Franz Pernkopf,
 and Holger Fröning*

Auxiliary Guided Autoregressive Variational Autoencoders 443
Thomas Lucas and Jakob Verbeek

Cooperative Multi-agent Policy Gradient 459
*Guillaume Bono, Jilles Steeve Dibangoye, Laëtitia Matignon,
 Florian Pereyron, and Olivier Simonin*

Parametric t-Distributed Stochastic Exemplar-Centered Embedding 477
Martin Renqiang Min, Hongyu Guo, and Dinghan Shen

Joint Autoencoders: A Flexible Meta-learning Framework 494
Baruch Epstein, Ron Meir, and Tomer Michaeli

Privacy Preserving Synthetic Data Release Using Deep Learning 510
*Nazmiye Ceren Abay, Yan Zhou, Murat Kantarcioglu,
 Bhavani Thuraisingham, and Latanya Sweeney*

On Finer Control of Information Flow in LSTMs 527
Hang Gao and Tim Oates

MaxGain: Regularisation of Neural Networks by Constraining
 Activation Magnitudes 541
Henry Gouk, Bernhard Pfahringer, Eibe Frank, and Michael J. Cree

Ontology Alignment Based on Word Embedding and Random Forest Classification	557
<i>Ikechukwu Nkisi-Orji, Nirmalie Wiratunga, Stewart Massie, Kit-Ying Hui, and Rachel Heaven</i>	
Domain Adaption in One-Shot Learning	573
<i>Nanqing Dong and Eric P. Xing</i>	
Ensemble Methods	
Axiomatic Characterization of AdaBoost and the Multiplicative Weight Update Procedure.	591
<i>Ibrahim Alabdulmohsin</i>	
Modular Dimensionality Reduction	605
<i>Henry W. J. Reeve, Tingting Mu, and Gavin Brown</i>	
Constructive Aggregation and Its Application to Forecasting with Dynamic Ensembles.	620
<i>Vitor Cerqueira, Fabio Pinto, Luis Torgo, Carlos Soares, and Nuno Moniz</i>	
MetaBags: Bagged Meta-Decision Trees for Regression.	637
<i>Jihed Khiari, Luis Moreira-Matias, Ammar Shaker, Bernard Ženko, and Sašo Džeroski</i>	
Evaluation	
Visualizing the Feature Importance for Black Box Models	655
<i>Giuseppe Casalicchio, Christoph Molnar, and Bernd Bischl</i>	
Efficient Estimation of AUC in a Sliding Window	671
<i>Nikolaj Tatti</i>	
Controlling and Visualizing the Precision-Recall Tradeoff for External Performance Indices.	687
<i>Blaise Hanczar and Mohamed Nadif</i>	
Evaluation Procedures for Forecasting with Spatio-Temporal Data.	703
<i>Mariana Oliveira, Luís Torgo, and Vítor Santos Costa</i>	
A Blended Metric for Multi-label Optimisation and Evaluation	719
<i>Laurence A. F. Park and Jesse Read</i>	
Author Index	735

Contents – Part II

Graphs

Temporally Evolving Community Detection and Prediction in Content-Centric Networks	3
<i>Ana Paula Appel, Renato L. F. Cunha, Charu C. Aggarwal, and Marcela Megumi Terakado</i>	
Local Topological Data Analysis to Uncover the Global Structure of Data Approaching Graph-Structured Topologies	19
<i>Robin Vandaele, Tijl De Bie, and Yvan Saeys</i>	
Similarity Modeling on Heterogeneous Networks via Automatic Path Discovery	37
<i>Carl Yang, Mengxiong Liu, Frank He, Xikun Zhang, Jian Peng, and Jiawei Han</i>	
Dynamic Hierarchies in Temporal Directed Networks	55
<i>Nikolaj Tatti</i>	
Risk-Averse Matchings over Uncertain Graph Databases	71
<i>Charalampos E. Tsourakakis, Shreyas Sekar, Johnson Lam, and Liu Yang</i>	
Discovering Urban Travel Demands Through Dynamic Zone Correlation in Location-Based Social Networks	88
<i>Wangsu Hu, Zijun Yao, Sen Yang, Shuhong Chen, and Peter J. Jin</i>	
Social-Affiliation Networks: Patterns and the SOAR Model	105
<i>Dhivya Eswaran, Reihaneh Rabbany, Artur W. Dubrawski, and Christos Faloutsos</i>	
ONE-M: Modeling the Co-evolution of Opinions and Network Connections	122
<i>Aastha Nigam, Kijung Shin, Ashwin Bahulkar, Bryan Hooi, David Hachen, Boleslaw K. Szymanski, Christos Faloutsos, and Nitesh V. Chawla</i>	
Think Before You Discard: Accurate Triangle Counting in Graph Streams with Deletions	141
<i>Kijung Shin, Jisu Kim, Bryan Hooi, and Christos Faloutsos</i>	

Semi-supervised Blockmodelling with Pairwise Guidance.	158
<i>Mohadeseh Ganji, Jeffrey Chan, Peter J. Stuckey, James Bailey, Christopher Leckie, Kotagiri Ramamohanarao, and Laurence Park</i>	

Kernel Methods

Large-Scale Nonlinear Variable Selection via Kernel Random Features	177
<i>Magda Gregorová, Jason Ramapuram, Alexandros Kalousis, and Stéphane Marchand-Maillet</i>	

Fast and Provably Effective Multi-view Classification with Landmark-Based SVM.	193
<i>Valentina Zantedeschi, Rémi Emonet, and Marc Sebban</i>	

Nyström-SGD: Fast Learning of Kernel-Classifiers with Conditioned Stochastic Gradient Descent	209
<i>Lukas Pfahler and Katharina Morik</i>	

Learning Paradigms

Hyperparameter Learning for Conditional Kernel Mean Embeddings with Rademacher Complexity Bounds	227
<i>Kelvin Hsu, Richard Nock, and Fabio Ramos</i>	

Deep Learning Architecture Search by Neuro-Cell-Based Evolution with Function-Preserving Mutations.	243
<i>Martin Wistuba</i>	

VC-Dimension Based Generalization Bounds for Relational Learning	259
<i>Ondřej Kuželka, Yuyi Wang, and Steven Schockaert</i>	

Robust Super-Level Set Estimation Using Gaussian Processes	276
<i>Andrea Zanette, Junzi Zhang, and Mykel J. Kochenderfer</i>	

Scalable Nonlinear AUC Maximization Methods.	292
<i>Majdi Khalid, Indrakshi Ray, and Hamidreza Chitsaz</i>	

Matrix and Tensor Analysis

Lambert Matrix Factorization	311
<i>Arto Klami, Jarkko Lagus, and Joseph Sakaya</i>	

Identifying and Alleviating Concept Drift in Streaming Tensor Decomposition	327
<i>Ravdeep Pasricha, Ekta Gujral, and Evangelos E. Papalexakis</i>	

<p>MASAGA: A Linearly-Convergent Stochastic First-Order Method for Optimization on Manifolds</p> <p><i>Reza Babanezhad, Issam H. Laradji, Alireza Shafaei, and Mark Schmidt</i></p>	<p>344</p>
<p>Block CUR: Decomposing Matrices Using Groups of Columns</p> <p><i>Urvashi Oswal, Swayambhoo Jain, Kevin S. Xu, and Brian Eriksson</i></p>	<p>360</p>
<p>Online and Active Learning</p>	
<p>SpectralLeader: Online Spectral Learning for Single Topic Models</p> <p><i>Tong Yu, Branislav Kveton, Zheng Wen, Hung Bui, and Ole J. Mengshoel</i></p>	<p>379</p>
<p>Online Learning of Weighted Relational Rules for Complex Event Recognition</p> <p><i>Nikos Katzouris, Evangelos Michelioudakis, Alexander Artikis, and Georgios Paliouras</i></p>	<p>396</p>
<p>Toward Interpretable Deep Reinforcement Learning with Linear Model U-Trees</p> <p><i>Guiliang Liu, Oliver Schulte, Wang Zhu, and Qingcan Li</i></p>	<p>414</p>
<p>Online Feature Selection by Adaptive Sub-gradient Methods</p> <p><i>Tingting Zhai, Hao Wang, Frédéric Koriche, and Yang Gao</i></p>	<p>430</p>
<p>Frame-Based Optimal Design</p> <p><i>Sebastian Mair, Yannick Rudolph, Vanessa Closius, and Ulf Brefeld</i></p>	<p>447</p>
<p>Hierarchical Active Learning with Proportion Feedback on Regions</p> <p><i>Zhipeng Luo and Milos Hauskrecht</i></p>	<p>464</p>
<p>Pattern and Sequence Mining</p>	
<p>An Efficient Algorithm for Computing Entropic Measures of Feature Subsets</p> <p><i>Frédéric Pennerath</i></p>	<p>483</p>
<p>Anytime Subgroup Discovery in Numerical Domains with Guarantees</p> <p><i>Aimene Belfodil, Adnene Belfodil, and Mehdi Kaytoue</i></p>	<p>500</p>
<p>Discovering Spatio-Temporal Latent Influence in Geographical Attention Dynamics</p> <p><i>Minoru Higuchi, Kanji Matsutani, Masahito Kumano, and Masahiro Kimura</i></p>	<p>517</p>

Mining Periodic Patterns with a MDL Criterion	535
<i>Esther Galbrun, Peggy Cellier, Nikolaj Tatti, Alexandre Termier, and Bruno Crémilleux</i>	
Revisiting Conditional Functional Dependency Discovery: Splitting the “C” from the “FD”	552
<i>Joeri Rammelaere and Floris Geerts</i>	
Sqn2Vec: Learning Sequence Representation via Sequential Patterns with a Gap Constraint	569
<i>Dang Nguyen, Wei Luo, Tu Dinh Nguyen, Svetha Venkatesh, and Dinh Phung</i>	
Mining Tree Patterns with Partially Injective Homomorphisms	585
<i>Till Hendrik Schulz, Tamás Horváth, Pascal Welke, and Stefan Wrobel</i>	
Probabilistic Models and Statistical Methods	
Variational Bayes for Mixture Models with Censored Data.	605
<i>Masahiro Kohjima, Tatsushi Matsubayashi, and Hiroyuki Toda</i>	
Exploration Enhanced Expected Improvement for Bayesian Optimization.	621
<i>Julian Berk, Vu Nguyen, Sunil Gupta, Santu Rana, and Svetha Venkatesh</i>	
A Left-to-Right Algorithm for Likelihood Estimation in Gamma-Poisson Factor Analysis	638
<i>Joan Capdevila, Jesús Cerquides, Jordi Torres, François Petitjean, and Wray Buntine</i>	
Causal Inference on Multivariate and Mixed-Type Data	655
<i>Alexander Marx and Jilles Vreeken</i>	
Recommender Systems	
POLAR: Attention-Based CNN for One-Shot Personalized Article Recommendation	675
<i>Zhengxiao Du, Jie Tang, and Yuhui Ding</i>	
Learning Multi-granularity Dynamic Network Representations for Social Recommendation	691
<i>Peng Liu, Lemei Zhang, and Jon Atle Gulla</i>	
GeoDCF: Deep Collaborative Filtering with Multifaceted Contextual Information in Location-Based Social Networks	709
<i>Dimitrios Rafailidis and Fabio Crestani</i>	

Personalized Thread Recommendation for MOOC Discussion Forums 725
*Andrew S. Lan, Jonathan C. Spencer, Ziqi Chen,
 Christopher G. Brinton, and Mung Chiang*

Inferring Continuous Latent Preference on Transition Intervals
 for Next Point-of-Interest Recommendation 741
Jing He, Xin Li, Lejian Liao, and Mingzhong Wang

Transfer Learning

Feature Selection for Unsupervised Domain Adaptation
 Using Optimal Transport 759
Leo Gautheron, Ievgen Redko, and Carole Lartizien

Web-Induced Heterogeneous Transfer Learning with Sample Selection 777
Sanatan Sukhija and Narayanan C. Krishnan

Towards More Reliable Transfer Learning 794
Zirui Wang and Jaime Carbonell

Differentially Private Hypothesis Transfer Learning 811
Yang Wang, Quanquan Gu, and Donald Brown

Information-Theoretic Transfer Learning Framework
 for Bayesian Optimisation 827
Anil Ramachandran, Sunil Gupta, Santu Rana, and Svetha Venkatesh

A Unified Framework for Domain Adaptation Using Metric
 Learning on Manifolds. 843
Sridhar Mahadevan, Bamdev Mishra, and Shalini Ghosh

Author Index 861

Contents – Part III

ADS Data Science Applications

Neural Article Pair Modeling for Wikipedia Sub-article Matching	3
<i>Muhao Chen, Changping Meng, Gang Huang, and Carlo Zaniolo</i>	
LinNet: Probabilistic Lineup Evaluation Through Network Embedding	20
<i>Konstantinos Pelechrinis</i>	
Improving Emotion Detection with Sub-clip Boosting	37
<i>Ermal Toto, Brendan J. Foley, and Elke A. Rundensteiner</i>	
Machine Learning for Targeted Assimilation of Satellite Data.	53
<i>Yu-Ju Lee, David Hall, Jebb Stewart, and Mark Govett</i>	
From Empirical Analysis to Public Policy: Evaluating Housing Systems for Homeless Youth	69
<i>Hau Chan, Eric Rice, Phebe Vayanos, Milind Tambe, and Matthew Morton</i>	
Discovering Groups of Signals in In-Vehicle Network Traces for Redundancy Detection and Functional Grouping	86
<i>Artur Mrowca, Barbara Moser, and Stephan Gunnemann</i>	

ADS E-commerce

SPEEDING Up the Metabolism in E-commerce by Reinforcement Mechanism DESIGN	105
<i>Hua-Lin He, Chun-Xiang Pan, Qing Da, and An-Xiang Zeng</i>	
Discovering Bayesian Market Views for Intelligent Asset Allocation	120
<i>Frank Z. Xing, Erik Cambria, Lorenzo Malandri, and Carlo Vercellis</i>	
Intent-Aware Audience Targeting for Ride-Hailing Service.	136
<i>Yuan Xia, Jingbo Zhou, Jingjia Cao, Yanyan Li, Fei Gao, Kun Liu, Haishan Wu, and Hui Xiong</i>	
A Recurrent Neural Network Survival Model: Predicting Web User Return Time	152
<i>Georg L. Grob, Ângelo Cardoso, C. H. Bryan Liu, Duncan A. Little, and Benjamin Paul Chamberlain</i>	
Implicit Linking of Food Entities in Social Media.	169
<i>Wen-Haw Chong and Ee-Peng Lim</i>	

A Practical Deep Online Ranking System in E-commerce Recommendation	186
<i>Yan Yan, Zitao Liu, Meng Zhao, Wentao Guo, Weipeng P. Yan, and Yongjun Bao</i>	

ADS Engineering and Design

Helping Your Docker Images to Spread Based on Explainable Models	205
<i>Riccardo Guidotti, Jacopo Soldani, Davide Neri, Antonio Brogi, and Dino Pedreschi</i>	

ST-DenNetFus: A New Deep Learning Approach for Network Demand Prediction	222
<i>Haytham Assem, Bora Caglayan, Teodora Sandra Buda, and Declan O’Sullivan</i>	

On Optimizing Operational Efficiency in Storage Systems via Deep Reinforcement Learning	238
<i>Sunil Srinivasa, Girish Kathalagiri, Julu Subramanyam Varanasi, Luis Carlos Quintela, Mohamad Charafeddine, and Chi-Hoon Lee</i>	

Automating Layout Synthesis with Constructive Preference Elicitation	254
<i>Luca Erculiani, Paolo Dragone, Stefano Teso, and Andrea Passerini</i>	

Configuration of Industrial Automation Solutions Using Multi-relational Recommender Systems	271
<i>Marcel Hildebrandt, Swathi Shyam Sunder, Serghei Mogoreanu, Ingo Thon, Volker Tresp, and Thomas Runkler</i>	

Learning Cheap and Novel Flight Itineraries	288
<i>Dmytro Karamshuk and David Matthews</i>	

Towards Resource-Efficient Classifiers for Always-On Monitoring	305
<i>Jonas Vlasselaer, Wannes Meert, and Marian Verhelst</i>	

ADS Financial/Security

Uncertainty Modelling in Deep Networks: Forecasting Short and Noisy Series	325
<i>Axel Brando, Jose A. Rodríguez-Serrano, Mauricio Ciprian, Roberto Maestre, and Jordi Vitrià</i>	

Using Reinforcement Learning to Conceal Honey-pot Functionality	341
<i>Seamus Dowling, Michael Schukat, and Enda Barrett</i>	

Flexible Inference for Cyberbully Incident Detection	356
<i>Haoti Zhong, David J. Miller, and Anna Squicciarini</i>	

Solving the False Positives Problem in Fraud Prediction Using Automated Feature Engineering	372
<i>Roy Wedge, James Max Kanter, Kalyan Veeramachaneni, Santiago Moral Rubio, and Sergio Iglesias Perez</i>	
Learning Tensor-Based Representations from Brain-Computer Interface Data for Cybersecurity	389
<i>Md. Lutfor Rahman, Sharmistha Bardhan, Ajaya Neupane, Evangelos Papalexakis, and Chengyu Song</i>	
ADS Health	
Can We Assess Mental Health Through Social Media and Smart Devices? Addressing Bias in Methodology and Evaluation	407
<i>Adam Tsakalidis, Maria Liakata, Theo Damoulas, and Alexandra I. Cristea</i>	
AMIE: Automatic Monitoring of Indoor Exercises	424
<i>Tom Decroos, Kurt Schütte, Tim Op De Beéck, Benedicte Vanwanseele, and Jesse Davis</i>	
Rough Set Theory as a Data Mining Technique: A Case Study in Epidemiology and Cancer Incidence Prediction	440
<i>Zaineb Chelly Dagdia, Christine Zarges, Benjamin Schannes, Martin Micallef, Lino Galiana, Benoît Rolland, Olivier de Fresnoye, and Mehdi Benchoufi</i>	
Bayesian Best-Arm Identification for Selecting Influenza Mitigation Strategies	456
<i>Pieter J. K. Libin, Timothy Verstraeten, Diederik M. Roijers, Jelena Grujic, Kristof Theys, Philippe Lemey, and Ann Nowé</i>	
Hypotensive Episode Prediction in ICUs via Observation Window Splitting	472
<i>Elad Tsur, Mark Last, Victor F. Garcia, Raphael Udassin, Moti Klein, and Evgeni Brotfain</i>	
Equipment Health Indicator Learning Using Deep Reinforcement Learning	488
<i>Chi Zhang, Chetan Gupta, Ahmed Farahat, Kosta Ristovski, and Dipanjan Ghosh</i>	
ADS Sensing/Positioning	
PBE: Driver Behavior Assessment Beyond Trajectory Profiling	507
<i>Bing He, Xiaolin Chen, Dian Zhang, Siyuan Liu, Dawei Han, and Lionel M. Ni</i>	

Accurate WiFi-Based Indoor Positioning with Continuous Location Sampling	524
<i>J. E. van Engelen, J. J. van Lier, F. W. Takes, and H. Trautmann</i>	
Human Activity Recognition with Convolutional Neural Networks	541
<i>Antonio Bevilacqua, Kyle MacDonald, Aamina Rangarej, Venessa Widjaya, Brian Caulfield, and Tahar Kechadi</i>	
Urban Sensing for Anomalous Event Detection:: Distinguishing Between Legitimate Traffic Changes and Abnormal Traffic Variability	553
<i>Masoomeh Zamani, Mengyi He, Masud Moshtaghi, Zahra Ghafouri, Christopher Leckie, James C. Bezdek, and Kotagiri Ramamohanarao</i>	
Combining Bayesian Inference and Clustering for Transport Mode Detection from Sparse and Noisy Geolocation Data.	569
<i>Danya Bachir, Ghazaleh Khodabandelou, Vincent Gauthier, Mounim El Yacoubi, and Eric Vachon</i>	
CentroidNet: A Deep Neural Network for Joint Object Localization and Counting	585
<i>K. Dijkstra, J. van de Loosdrecht, L. R. B. Schomaker, and M. A. Wiering</i>	
Deep Modular Multimodal Fusion on Multiple Sensors for Volcano Activity Recognition	602
<i>Hiep V. Le, Tsuyoshi Murata, and Masato Iguchi</i>	
Nectar Track	
Matrix Completion Under Interval Uncertainty: Highlights	621
<i>Jakub Marecek, Peter Richtarik, and Martin Takac</i>	
A Two-Step Approach for the Prediction of Mood Levels Based on Diary Data	626
<i>Vincent Bremer, Dennis Becker, Tobias Genz, Burkhardt Funk, and Dirk Lehr</i>	
Best Practices to Train Deep Models on Imbalanced Datasets—A Case Study on Animal Detection in Aerial Imagery	630
<i>Benjamin Kellenberger, Diego Marcos, and Devis Tuia</i>	
Deep Query Ranking for Question Answering over Knowledge Bases	635
<i>Hamid Zafar, Giulio Napolitano, and Jens Lehmann</i>	
Machine Learning Approaches to Hybrid Music Recommender Systems	639
<i>Andreu Vall and Gerhard Widmer</i>	

Demo Track

IDEA: An Interactive Dialogue Translation Demo System	
Using Furhat Robots	645
<i>Jinhua Du, Darragh Blake, Longyue Wang, Clare Conran, Declan Mckibben, and Andy Way</i>	
RAPID: Real-time Analytics Platform for Interactive Data Mining	649
<i>Kwan Hui Lim, Sachini Jayasekara, Shanika Karunasekera, Aaron Harwood, Lucia Falzon, John Dunn, and Glenn Burgess</i>	
Interactive Time Series Clustering with COBRAS^{TS}	654
<i>Toon Van Craenendonck, Wannes Meert, Sebastijan Dumančić, and Hendrik Blockeel</i>	
pysubgroup: Easy-to-Use Subgroup Discovery in Python	658
<i>Florian Lemmerich and Martin Becker</i>	
An Advert Creation System for Next-Gen Publicity	663
<i>Atul Nautiyal, Killian McCabe, Murhaf Hossari, Soumyabrata Dev, Matthew Nicholson, Clare Conran, Declan McKibben, Jian Tang, Wei Xu, and François Pitié</i>	
VHI: Valve Health Identification for the Maintenance of Subsea Industrial Equipment	668
<i>M. Atif Qureshi, Luis Miralles-Pechuán, Jing Su, Jason Payne, and Ronan O'Malley</i>	
Tiler: Software for Human-Guided Data Exploration	672
<i>Andreas Henelius, Emilia Oikarinen, and Kai Puolamäki</i>	
ADAGIO: Interactive Experimentation with Adversarial Attack and Defense for Audio	677
<i>Nilaksh Das, Madhuri Shanbhogue, Shang-Tse Chen, Li Chen, Michael E. Kounavis, and Duen Horng Chau</i>	
ClaRe: Classification and Regression Tool for Multivariate Time Series	682
<i>Ricardo Cachucho, Stylianos Paraschiakos, Kaihua Liu, Benjamin van der Burgh, and Arno Knobbe</i>	
Industrial Memories: Exploring the Findings of Government Inquiries with Neural Word Embedding and Machine Learning	687
<i>Susan Leavy, Emilie Pine, and Mark T. Keane</i>	
Monitoring Emergency First Responders' Activities via Gradient Boosting and Inertial Sensor Data	691
<i>Sebastian Scheurer, Salvatore Tedesco, Óscar Manzano, Kenneth N. Brown, and Brendan O'Flynn</i>	

**Visualizing Multi-document Semantics via Open Domain
Information Extraction** 695
*Yongpan Sheng, Zenglin Xu, Yafang Wang, Xiangyu Zhang, Jia Jia,
Zhonghui You, and Gerard de Melo*

Author Index 701