Lecture Notes in Artificial Intelligence 11051

Subseries of Lecture Notes in Computer Science

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

Springer

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

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

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

Applied Data Science Track Program Committee

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

Nectar Track Program Committee

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

Combatting 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

Applications

Classification

Clustering and Unsupervised Learning

Contents – Part II

Graphs

Online and Active Learning

Pattern and Sequence Mining

Transfer Learning

Contents – Part III

Demo Track