

Preface

We are delighted to introduce the proceedings of the 2017 edition of the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD 2017). This year the conference was held in Skopje, Macedonia, during September 18–22, 2017. 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 as well as innovative application domains. This event is the premier European machine learning and data mining conference and builds upon a very successful series of 18 ECML, ten PKDD (until 2007, when they merged), and nine ECML PKDD (from 2008). Therefore, this was the tenth edition of ECML PKDD as a single conference.

The scientific program was very rich 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, featuring regular conference papers, published in the proceedings
- The journal track, featuring papers that satisfy the quality criteria of journal papers and at the same time lend themselves to conference talks (these papers are published separately in the journals *Machine Learning* and *Data Mining and Knowledge Discovery*)
- The applied data science track (formerly industrial track), aiming to bring together participants from academia, industry, government, and non-governmental organizations in a venue that highlights practical and real-world studies of machine learning, knowledge discovery, and data mining
- The demo track, presenting innovative prototype implementations or mature systems that use machine learning techniques and knowledge discovery processes in a real setting
- The nectar track, offering conference attendees a compact overview of recent scientific advances at the frontier of machine learning and data mining with other disciplines, as published in related conferences and journals.

In addition, the PhD Forum provided a friendly environment for junior PhD students to exchange ideas and experiences with peers in an interactive atmosphere and to get constructive feedback from senior researchers. This year we also introduced the EU Projects Forum with the purpose of disseminating EU projects and their results to the targeted scientific audience of the conference participants. Moreover, two discovery challenges, six tutorials, and 17 co-located workshops were organized on related research topics.

Following the successful experience of last year, we stimulated the practices of reproducible research (RR). Authors were encouraged to adhere to such practices by making available data and software tools for reproducing the results reported in their

papers. In total, 57% of the accepted papers in the conference have accompanying software and/or data and are flagged as RR papers on the conference website. In the proceedings, each RR paper provides links to such additional material. This year, we took a further step toward data publishing and facilitated the authors in making data and software available on a public repository in a dedicated area branded with the conference name (e.g., figshare, where there is an ECML PKDD area).

The response to our call for papers was very good. We received 364 papers for the main conference track, of which 104 were accepted, yielding an acceptance rate of about 28%. This allowed us to define a very rich program with 101 presentations in the main conference track: The remaining three accepted papers were not presented at the conference and are not included in the proceedings. The program also included six plenary keynotes by invited speakers: Inderjit Dhillon (University of Texas at Austin and Amazon), Alex Graves (Google DeepMind), Frank Hutter (University of Freiburg), Pierre-Philippe Mathieu (ESA/ESRIN), John Quackenbush (Dana-Farber Cancer Institute and Harvard TH Chan School of Public Health), and Cordelia Schmid (Inria).

This year, ECML PKDD attracted over 600 participants from 49 countries. It also attracted substantial attention from industry and end users, both through sponsorship and submission/participation at the applied data science track. This is confirmation that the ECML PKDD community is healthy, strong, and continuously growing.

Many people worked hard together as a superb dedicated team to ensure the successful organization of this conference: the general chairs (Michelangelo Ceci and Sašo Džeroski), the PC chairs (Michelangelo Ceci, Jaakko Hollmén, Ljupčo Todorovski, Celine Vens), the journal track chairs (Kurt Driessens, Dragi Kocev, Marko Robnik-Šikonja, Myra Spiliopoulou), the applied data science chairs (Yasemin Altun, Kamalika Das, Taneli Mielikäinen), the nectar track chairs (Donato Malerba, Jerzy Stefanowski), the demonstration track chairs (Jesse Read, Marinka Žitnik), the workshops and tutorials chairs (Nathalie Japkowicz, Panče Panov), the EU projects forum chairs (Petra Kralj Novak, Nada Lavrač), the PhD forum chairs (Tomislav Šmuc, Bernard Ženko), the Awards Committee (Peter Flach, Rosa Meo, Indrè Žliobaitė), the discovery challenge chair (Dino Ienco), the production and public relations chairs (Dragi Kocev, Nikola Simidjievski), the local organizers (Ivica Dimitrovski, Tina Anžič, Mili Bauer, Gjorgji Madjarov), the sponsorship chairs (Albert Bifet, Panče Panov), the proceedings chairs (Jurica Levatić, Gianvito Pio), the area chairs (listed in this book), the Program Committee members (listed in this book), and members of the MLJ and DMKD Guest Editorial Boards. They made tremendous effort in guaranteeing the quality of the reviewing process and, therefore, the scientific quality of the conference, which is certainly beneficial for the whole community. Our sincere thanks go to all of them.

We would also like to thank the Cankarjev Dom congress agency and the student volunteers. Thanks to Springer for their continuous support and Microsoft for allowing us to use their CMT software for conference management. Special thanks to the sponsors (Deutsche Post DHL Group, Google, AGT, ASML, Deloitte, NEC, Siemens, Cambridge University Press, *IEEE/CAA Journal of Automatica Sinica*, Springer, IBM Research, *Data Mining and Knowledge Discovery*, *Machine Learning*, EurAi, and

GrabIT) and the European project MAESTRA (ICT-2013-612944), as well as the ECML PKDD Steering Committee (for their suggestions and advice). Finally, we would like to thank the organizing institutions: the Jožef Stefan Institute (Slovenia), the Ss. Cyril and Methodius University in Skopje (Macedonia), and the University of Bari Aldo Moro (Italy).

September 2017

Michelangelo Ceci
Jaakko Hollmén
Ljupčo Todorovski
Celine Vens
Sašo Džeroski

Machine Learning and Knowledge Discovery in
Databases

European Conference, ECML PKDD 2017, Skopje,
Macedonia, September 18–22, 2017, Proceedings, Part
I

Ceci, M.; Hollmén, J.; Todorovski, L.; Vens, C.; Džeroski,
S. (Eds.)

2017, LXIII, 852 p. 245 illus., Softcover

ISBN: 978-3-319-71248-2