

## Guest editors' introduction: special issue of selected papers from ECML-PKDD 2012

Tijl De Bie · Peter A. Flach

Received: 30 May 2013 / Accepted: 31 May 2013 / Published online: 6 June 2013  
© The Author(s) 2013

The 2012 instalment of the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML-PKDD) was held in Bristol, United Kingdom during the week of 24–28 September 2012. ECML-PKDD is a highly successful and selective international conference series which was first organised in its present form in 2001 in Freiburg, Germany, when it joined together the hitherto separate ECML and PKDD conferences. Since then, the two strands of machine learning and data mining have been increasingly integrated in the joint conference, and today it is well-known as the only top-tier international conference which exploits the synergy between these two exciting fields.

The format of the 2012 conference followed the tried-and-tested format of previous editions, with workshops and tutorials on Monday and Friday; research papers in parallel tracks on Tuesday, Wednesday and Thursday; and plenary keynote talks on each of the five conference days. The 105 research papers were carefully selected from 443 submitted papers (acceptance rate 23.7 %) on the basis of reviews by 275 Program Committee members and 36 Area Chairs, assisted by 161 additional reviewers. The proceedings were published in two volumes of Springer's Lecture Notes in Artificial Intelligence series (Flach et al. 2012a, 2012b).

After the conference, the authors of the nine most highly evaluated machine learning papers were invited to submit a significantly extended version of their paper to this special issue. Submitted papers were assessed by three reviewers, some of whom had reviewed the original conference submission and some of whom had not seen the paper before. After several rounds of reviewing and revision, seven of the submitted papers were selected for inclusion. The selected papers address a range of learning tasks and applications and indicate the breadth and sophistication of current-day machine learning research. We will now briefly describe the contribution that each paper makes; the papers are ordered alphabetically on title.

---

T. De Bie · P.A. Flach (✉)  
Intelligent Systems Laboratory, University of Bristol, Bristol, UK  
e-mail: [Peter.Flach@bristol.ac.uk](mailto:Peter.Flach@bristol.ac.uk)

The paper ‘A reinforcement learning approach to autonomous decision-making in smart electricity markets’ by Markus Peters, Wolfgang Ketter, Maytal Saar-Tsechansky and John Collins proposes a class of autonomous broker agents for retail electricity trading that can operate in a wide range of Smart Electricity Markets, and that are capable of deriving long-term, profit-maximising policies. The authors show how feature selection and regularisation can be leveraged to automatically optimise brokers for particular market conditions, and evaluate performance using real-world energy market data.

‘Active evaluation of ranking functions based on graded relevance’ by Christoph Sawade, Steffen Bickel, Timo von Oertzen, Tobias Scheffer and Niels Landwehr studies the problem of estimating ranking performance under time-varying query distributions and item relevance. Estimates are based on a set of most informative test queries selected by an active sampling distribution. The authors derive cost-optimal sampling distributions for two commonly used performance measures and experiments on web search engine data illustrate significant reductions in labelling costs. The conference version of this paper received the best paper award in the machine learning area sponsored by *Machine Learning*.

The paper ‘Beam search algorithms for multilabel learning’ by Abhishek Kumar, Shankar Vembu, Aditya Krishna Menon and Charles Elkan extends the popular probabilistic classifier chain method that predicts multiple correlated labels. Using the classical AI technique of beam search the proposed method makes prediction tractable and integrates the search for a label order into the learning phase.

‘Exploiting symmetries for scaling loopy belief propagation and relational training’ by Babak Ahmadi, Kristian Kersting, Martin Mladenov and Sriraam Natarajan studies the problem of lifting propositional belief propagation techniques to relational models. The authors demonstrate that their method is amenable to MapReduce parallelisation and hence scaleable, achieving an order of magnitude speed improvement in online training without loss of quality in the solution.

The next paper ‘Geometry preserving multi-task metric learning’ by Peipei Yang, Kaizhu Huang and Cheng-Lin Liu considers the problem of simultaneously learning multiple metrics from correlated tasks. Using von Neumann divergence, the authors are able to preserve the geometry of the data leading to more appropriate propagation of side-information between tasks.

‘Learning policies for battery usage optimisation in electric vehicles’ by Stefano Ermon, Yexiang Xue, Carla Gomes and Bart Selman addresses the problem of optimising real-time energy management of electric vehicles exploiting multiple batteries. The problem is formalised as a stochastic planning problem, and the authors propose a novel solution based on a combination of optimisation, machine learning and data-mining techniques. Experiments on various data sets of commuter trips demonstrate that the policy outperforms algorithms that were previously proposed as part of an open algorithmic challenge.

Finally, ‘Massively parallel feature selection: an approach based on variance preservation’ by Zheng Zhao, Ruiwen Zhang, James Cox, David Duling and Warren Sarle presents a large-scale feature selection algorithm that favours features that explain data variance. By defining the preserving target in different ways, the algorithm can achieve both supervised and unsupervised feature selection and supports regression as well as classification. The authors demonstrate the performance of the algorithm on data sets of up to 10,000 features.

This special issue wouldn’t have been possible without the cooperation and help of many people. We thank the ECML-PKDD 2012 area chairs for providing a long-list of candidate papers; the reviewers for their insightful comments which often significantly improved the papers; and the authors for choosing *Machine Learning* to publish their work.

## References

- Flach, P. A., De Bie, T., & Cristianini, N. (Eds.) (2012a). *Lecture notes in computer science: Vol. 7523. Machine learning and knowledge discovery in databases—European conference, ECML PKDD 2012. Proceedings, part I*. Berlin: Springer. Bristol, UK, 24–28 September 2012.
- Flach, P. A., De Bie, T., & Cristianini, N. (Eds.) (2012b). *Lecture notes in computer science: Vol. 7524. Machine learning and knowledge discovery in databases—European conference, ECML PKDD 2012. Proceedings, part II*. Berlin: Springer. Bristol, UK, 24–28 September 2012.