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

Tijl De Bie · Peter A. Flach

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 5 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,b).

After the conference, the authors of the four most highly evaluated data mining and knowledge discovery papers were invited to submit a significantly extended version of their paper to this special section. 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, two of the submitted papers were selected for inclusion. 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: tijl.debie@gmail.com

The paper 'An efficiently computable subgraph pattern support measure: counting independent observations' by Yuyi Wang, Jan Ramon, and Thomas Fannes introduces a new measure to quantify the support of a subgraph pattern in a given database graph. The authors prove that this measure is normalized and anti-monotonic, and that it can be computed efficiently using a sparse linear program, making it the first overlap-based subgraph support measure that is amenable to use on large networks. As a theoretical side-result, the authors also relate the proposed support measure to the statistical power of the sample of embeddings of the subgraph pattern.

The paper 'Nearly exact mining of frequent trees in large networks' by Ashraf M. Kibriya and Jan Ramon presents an algorithm to mine all frequent rooted trees smaller than a specified size in a single network. Exploiting recent insights from the theory of parameterized complexity, the authors achieved a delay complexity that is linear in the network size, making the algorithm applicable to very large networks.

This special section 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 *Data Mining and Knowledge Discovery* to publish their work.

References

Peter A. Flach, Tijl De Bie, and Nello Cristianini, eds (2012) Machine learning and knowledge discovery in databases—European conference, ECML PKDD 2012. Bristol, September 24–28. Proceedings, Part I, vol. 7523 of Lecture Notes in Computer Science. Springer, Berlin

Peter A. Flach, Tijl De Bie, and Nello Cristianini, eds (2012) Machine learning and knowledge discovery in databases—European conference, ECML PKDD 2012, Bristol, September 24–28, 2012. Proceedings, Part II, vol 7524 of Lecture Notes in Computer Science. Springer, Berlin