

ECML/PKDD01 Workshop

Integrating Aspects of Data Mining, Decision Support and Meta-learning (IDDM-2001)

Christophe Giraud-Carrier
ELCA Informatique SA, CH-1000 Lausanne 13, Switzerland

Nada Lavrač
J. Stefan Institute, 1000 Ljubljana, Slovenia

Steve Moyle
Oxford University Computing Laboratory, Oxford OX1 3QD, UK

Overall Aim

This workshop is aimed at researchers and practitioners in Data Mining, Decision Support, and Machine Learning. Participants will gain a better appreciation of the issues facing the application and deployment of Data Mining solutions in the real world. New ways of working together and combining results will be discussed, fostering further collaboration between participants' organisations. It is hoped that more people will work together more often and in more sensible ways as a result of this workshop.

This workshop continues in the tradition of previous related workshops, such as the ICML'97 Workshop on Machine Learning Applications in the Real World, the AAAI'98/ICML'98 Workshop on the Methodology of Applying Machine Learning, the ICML'99 Workshop on Advances in Meta-learning and Future Work, and the ECML'00 joint Familiarisation Workshop for the METAL, MiningMarts, Sol-Eu-Net and MLNET projects.

Motivation and Scope

The CRISP Data Mining methodology provides guidelines and a sequence of steps to be followed in the applied knowledge discovery process. Decision Support and Meta-learning are further valuable approaches to knowledge discovery. This workshop is aimed at exploring opportunities for integrating aspects of Data Mining, Decision Support, and Meta-learning in applied knowledge discovery, including extending the CRISP Data Mining methodology through contributions addressing the following issues:

Integration of different methods on the same problem

Novel ideas and reviews of existing approaches to combining results of classifiers, meta-learning, etc., with emphasis on model selection, model combination and all issues relevant to learning to learn (e.g., landmarking, performance prediction, knowledge transfer, data characterisation, meta-data collection and exploitation, standardised experimental set-ups/methods, etc.).

Integration of work of different groups/individuals on the same problem

Novel ideas and reviews of existing approaches to collaborative Data Mining. Usually, Data Mining tasks are solved by a single individual or group of individuals working jointly on a problem. However, with the Internet, Data Mining tasks could be solved through a collaboration of different groups of researchers at different sites. Different modes of collaboration should be explored (e.g., competitive vs. collaborative) and issues such as infrastructure and methods for supporting distant collaborative work (e.g., how to integrate new individuals/groups following the start/stop-any-time principle) should be addressed.

Integration of different problem solving approaches, particularly Data Mining and Decision Support

Data Mining has the potential of solving Decision Support problems, when previous Decision Support solutions have been recorded as data to be used for analysis with DM tools. On the other hand, Decision Support methodology usually results in a decision model, reflecting expert knowledge of decision makers. How can such expert knowledge be incorporated into problem solutions by Data Mining? Can it be used as background knowledge in relational Data Mining? Can such expert knowledge be induced automatically? Are there any systematic methodological means of combining the two approaches to problem solving?

Integration of results of models from different datasets

Data in standard Data Mining often has the form of a single relational table. What if data is stored in multiple relational tables? This topic considers relational Data Mining through combining the results of mining separate relational tables. A standard approach in ILP is to consider one table as the master data table, and all others as tables providing background knowledge. What if this is not natural? Would mining of individual tables and combining results be a better solution? Are there other approaches to this problem?

Integration by learning from successes and failures

Journal, conference and workshop papers seldom report on failures, yet reported failures are crucial for increasing awareness of which Data Mining approach/algorithm to use in which situation. Analyses of failures of individual approaches, as well as of combined approaches to Data Mining are welcome. These could also include analyses of successes and failures in challenge problems such as COIL, Sisyphus, PTE, etc.

Organisation

Programme Chairs

C. Giraud-Carrier, ELCA Informatique SA, Switzerland
N. Lavrac, Institute Jozef Stefan, Slovenia
S. Moyle, Oxford University, United Kingdom

Programme Committee

Marko Bohanec, Institute Jozef Stefan, Slovenia
Pavel Brazdil, LIACC, University of Porto, Portugal
Philip Chan, Florida Institute of Technology, USA
Saso Dzeroski, Institute Jozef Stefan, Slovenia
Peter Flach, University of Bristol, United Kingdom
Dragan Gamberger, Rudjer Boskovic Institute, Croatia
Marko Grobelnik, Institute Jozef Stefan, Slovenia
Alipio Jorge, LIACC, University of Porto, Portugal
Stan Matwin, University of Ottawa, Canada
Dunja Mladenic, Institute Jozef Stefan, Slovenia
Ljupco Todorovski, Institute Jozef Stefan, Slovenia
Maarten van Someren, University of Amsterdam, The Netherlands
Ricardo Vilalta, IBM T.J. Watson Research Center, USA
Takahira Yamaguchi, Shizuoka University, Japan

Acknowledgements

We gratefully acknowledge support from the European Commission's projects METAL and Sol-Eu-Net.

We also wish to thank the members of the programme committee for their timely and thorough reviews of the papers, as well as for their many constructive suggestions to the authors. Finally, we thank Branko Kavšek for setting up and managing Web facilities for the review process and the electronic publication of the proceedings. We hope you find the workshop's material and presentations stimulating.

C. Giraud-Carrier, N. Lavrač and S. Moyle
Freiburg, September 2001