Foreword

This volume contains the papers accepted to the 5th Asian Conference on Machine Learning (ACML 2013), following successful preceding conferences held in China, Japan, Taiwan and Singapore. The conference was held on 13-15 November 2013, at the Australian National University, Canberra, Australia. ACML aims at providing a leading international forum for researchers in machine learning and related fields to share their new ideas and achievements. Submissions from other than the Asia-Pacific regions were also highly encouraged.

The conference called for research papers reporting original investigation results, and received 103 paper submissions from 22 countries. Following the successful two round scheme of the previous year, we had an early deadline in May 2013 and a final deadline in July 2013. Some of the early round papers were revised and resubmitted. We eventually accepted 32 papers (31% acceptance rate) with 13 long presentations and 19 short presentations. Every paper was double-blind reviewed by at least 3 expert reviewers, and these were then used to form the opinion of 2 senior program committee members. To ensure quality, each reviewer had a quota of no more than 6 papers to review. In total, there were 26 senior program committee members and a further 100 expert reviewers who contributed to the high quality program at ACML. Their names are acknowledged in the following pages. Finally, the program co-chairs considered all the reviews and meta-reviews by senior program committee members to make the final decisions for the papers.

All accepted papers received both an oral and poster presentation, and are published in Volume 29 of the JMLR Workshop and Proceedings track. A special thanks goes to the Publications Chair Justin Domke for diligently organising the camera ready papers. Following the tradition of previous ACMLs, this year's conference is also single track. The submissions covered a broad range of topics, including theoretical analyses, algorithmic improvements in optimisation methods for machine learning, and applications to real world problems.

In addition to the submitted papers, we were pleased to have three keynotes from leading experts. Geoffrey Holmes talked about "Weka, MOA and Experiment Databases: Frameworks for Machine Learning", Chih-Jen Lin discussed "Recent advances in large linear classification" and Ralf Herbrich explained "From Theory to Real-World Applications: Technology Transfer in Practice". As in previous years, we had a day of tutorials and workshops prior to the main conference organised by the Tutorial and Workshops Chair Lexing Xie. The tutorials were "Distributed, Real-Time Bayesian Learning in Online Service" by Ralf Herbrich, "Cost-sensitive Classification: Algorithms and Advances [slides]" by Hsuan-Tien Lin, and "Recent Advances in Bayesian Methods" by Jun Zhu. In parallel to the tutorials, we had a workshop on "Theory and Practice in Machine Learning" organised by Peter Sunehag, Marcus Hutter, and Mark Reid. We thank all the speakers and organisers for putting together a fantastic program.

We gratefully acknowledge our financial sponsors: NICTA, the Air Force Office of Scientific Research USA, the Asian Office of Aerospace R&D USA, the Australian National University, Rulequest Research and Machine Learning Journal.

ACML Steering Committee Co-Chair Zhi-Hua Zhou provided valuable advise and support during the whole process, and Steering Committee Members Hiroshi Motoda, Masashi Sugiyama and Geoff Webb provided great help in selecting the best paper and runner-up paper. Wray Buntine and Bob Williamson, our General Chairs, made sure that ACML ran smoothly and Web Chair Xinhua Zhang was up all hours keeping the various bits of our website up to date. Our special thanks goes to Local Arrangements Chairs Sarika Abbott and Krystel Boland who made the day to day running of the conference possible.

Last but not least, a big thank you to all participants of ACML 2013 who made it such a great event!

November 2013

ICML 2013 Program Co-Chairs

Cheng Soon Ong
NICTA, Melbourne, Australia
cheng-soon.ong@nicta.com.au

Tu Bao Ho
School of Knowledge Science, JAIST, Japan
bao@jaist.ac.jp

Invited Talks

Title: Weka, MOA and Experiment Databases: Frameworks for Machine Learning

Presenter: Geoffrey Holmes

Abstract:

This talk is in three parts. The first deals with an aspect of the Weka project that has received little attention, namely the use of machine learning in agricultural applications. I will outline our experiences in this field and present an application development framework which is a direct result of this activity. In particular, one project has met one of the challenges proposed by Kiri Wagstaff at ICML 2012. Second, I will talk about our work in data stream mining with a focus on classification within the Massive Online Analysis framework MOA. After a quick overview of what is in MOA I will present two recent results that indicate a need for caution and a statement of what constitutes state-of-the-art in data stream classification for practitioners. Finally, I will present the idea of experiment databases, a framework for machine learning experimentation that saves effort and offers opportunities for meta learning and hypothesis generation.

Bio:

Geoff is currently Dean of the Faculty of Computing and Mathematical Sciences at the University of Waikato. He has been head of the machine learning group and has been involved in several open source projects over the last 20 years. He has made contributions in machine learning across several branches of the subject and has been active in finding ways to reward researchers for their efforts to produce open source software. In this regard he acts as an action editor for the branch of JMLR dedicated to open source software. He was part of the team that in 2005 won the SIGKDD Data Mining and Knowledge Discovery Service Award for Weka and regularly serves on senior PCs for KDD, ECMLPKDD and Discovery Science.

He obtained BSc and PhD degrees in Mathematics from Southampton University, UK in 1986. After time as a research assistant in Cambridge University he joined Waikato in 1987, after moving up the ranks, he was promoted to Professor in 2008.

Title: Recent Advances in Large Linear Classification

Presenter: Chih-Jen Lin

Abstract:

Linear classification is a useful tool in machine learning and data mining. For some data in a rich dimensional space, the prediction performance of linear classifiers has shown to be close to that of nonlinear classifiers such as kernel methods, but training and testing speed

is much faster. Recently, many research works have proposed efficient optimization methods to construct linear classifiers. We briefly discuss some of them that were considered in our development of the software LIBLINEAR. We then move to discuss some extensions of linear classification. In particular, linear classifiers can be useful to either directly or indirectly approximate kernel classifiers. I will show some real-word examples for which we try to achieve fast training/testing speed, while maintain competitive accuracy. Finally, future challenges of this research topic, in particular, aspects on big-data linear classification, will be discussed.

Bio:

Chih-Jen Lin is currently a distinguished professor at the Department of Computer Science, National Taiwan University. He obtained his B.S. degree from National Taiwan University in 1993 and Ph.D. degree from University of Michigan in 1998. His major research areas include machine learning, data mining, and numerical optimization. He is best known for his work on support vector machines (SVM) for data classification. His software LIBSVM is one of the most widely used and cited SVM packages. For his research work he has received many awards, including the ACM KDD 2010 and ACM RecSys 2013 best paper awards. He is an IEEE fellow and an ACM distinguished scientist for his contribution to machine learning algorithms and software design. More information about him can be found at http://www.csie.ntu.edu.tw/ cjlin.

Title: From Theory to Real-World Applications: Technology Transfer in Practice

Presenter: Ralf Herbrich

Abstract:

In this talk, I will talk about several of my experiences of technology transfer over the past 10 years at Microsoft, Facebook and Amazon. Technology transfer - the process of moving promising research topics into a level of maturity that's ready for bulk manufacturing or production - remains one of the largest mysteries and unsolved social and organizational problems. There are some surprising similarities in the various technology transfers of mathematical models and algorithms into products and I will attempt to group and discuss them. During the course of the talk, I will introduce graphical models and approximate message passing as well as their application to gamer ranking and computational online advertising.

Bio:

Ralf Herbrich is Director of Machine Learning at Amazon. From 2011 to 2012, he worked at Facebook leading the Unified Ranking and Allocation team building large-scale machine learning infrastructure for learning user-action-rate predictors that enabled unified value

experiences across Facebook products. From 2009 to 2011, he was Director of Microsoft's Future Social Experiences (FUSE) Lab UK working on the development of computational intelligence technologies on large online data collections. From 2006 to 2010, Ralf was co-leading the Applied Games and Online Services and Advertising group at Microsoft Research Cambridge which engaged in research at the intersection of machine learning and computer games and in the areas of online services, search and online advertising combining insights from machine learning, information retrieval, game theory, artificial intelligence and social network analysis. Ralf joined Microsoft Research in 2000 as a Postdoctoral researcher and Research Fellow of the Darwin College Cambridge. Prior to joining Microsoft, he obtained both a diploma degree in Computer Science in 1997 and a Ph.D. degree in Statistics in 2000 from Technical University of Berlin. Ralf has published over 80 papers and holds over 30 patents. His research interests include Bayesian inference and decision making, kernel methods, statistical learning theory, distributed systems and programming languages. Ralf is one of the inventors of the Drivatars$^{\text{TM}}$system in the Forza Motorsport series as well as the TrueSkill$^{\text{TM}}$ranking and matchmaking system in Xbox 360 Live. He also co-invented the adPredictor click-prediction technology launched in 2009 in Bing's online advertising system.

Tutorials

Title: Distributed, Real-Time Bayesian Learning in Online Service

Presenter: Ralf Herbrich

Abstract:

The last ten years have seen a tremendous growth in Internet-based online services such as search, advertising, gaming and social networking. Today, it is important to analyze large collections of user interaction data as a first step in building predictive models for these services as well as learn these models in real-time. One of the biggest challenges in this setting is scale: not only does the sheer scale of data necessitate parallel processing but it also necessitates distributed models; with hundreds of million active users on major online services such as Facebook, Twitter, Amazon or Google, any user-specific sets of features in a linear or non-linear model yields models of a size bigger than can be stored in a single system.

In this tutorial, I will give an introduction to distributed message passing, a theoretical framework that can deal both with the distributed inference and storage of models. After an overview of message passing, I will discuss and present recent advances in approximate message passing which allows to control the model size as the amount of training data grows. We will also review how distributed (approximate) message passing can be mapped to generalized distributed computing and how modeling constraints map on the system design. In the second part of the talk, I will give an overview of the application of these techniques to real-world learning systems, namely:

Gamer ranking and matchmaking in TrueSkill^{TM}and Halo 3 AdPredictor click-through rate learning and prediction in sponsored search User-action models in Facebook's information distribution and advertising pipeline

Bio:

Ralf Herbrich is Director of Machine Learning at Amazon. From 2011 to 2012, he worked at Facebook leading the Unified Ranking and Allocation team building large-scale machine learning infrastructure for learning user-action-rate predictors that enabled unified value experiences across Facebook products. From 2009 to 2011, he was Director of Microsoft's Future Social Experiences (FUSE) Lab UK working on the development of computational intelligence technologies on large online data collections. From 2006 to 2010, Ralf was co-leading the Applied Games and Online Services and Advertising group at Microsoft Research Cambridge which engaged in research at the intersection of machine learning and computer games and in the areas of online services, search and online advertising combining insights from machine learning, information retrieval, game theory, artificial intelligence and social network analysis. Ralf joined Microsoft Research in 2000 as a Postdoctoral researcher and Research Fellow of the Darwin College Cambridge. Prior to joining Microsoft, he obtained both a diploma degree in Computer Science in 1997 and a Ph.D. degree in Statistics in 2000 from Technical University of Berlin. Ralf has published over 80 papers and holds over 30 patents. His research interests include Bayesian inference and decision making, kernel methods, statistical learning theory, distributed systems and programming languages. Ralf is one of the inventors of the Drivatars^{TM}system in the Forza Motorsport series as well as

the TrueSkill^{TM}ranking and matchmaking system in Xbox 360 Live. He also co-invented the adPredictor click-prediction technology launched in 2009 in Bing's online advertising system.

Title: Cost-sensitive Classification: Algorithms and Advances

Presenter: Hsuan-Tien Lin

Abstract:

Classification is an important problem in machine learning. It can be used in a variety of applications, such as distinguishing apples, oranges, and bananas automatically. Traditionally, the regular classification problem aims at minimizing the number of future mis-prediction errors. Nevertheless, many real-world applications require varying costs for different types of mis-classification errors. For instance, a false-negative prediction for a spam classification system only takes the user an extra second to delete the email, while a false-positive prediction can mean a huge loss when the email actually carries important information; in bacteria classification, mis-classifying a Gram-positive species as a Gram-negative one leads to totally ineffective treatments and is hence more serious than mis-classifying a Gram-positive species as another Gram-positive one; when classifying a patient as healthy, cold-infected, or H1N1-infected, predicting an H1N1-infected patient as healthy is significantly more serious than predicting a predicting a healthy patient as H1N1-infected. Such a cost-sensitive classification problem can be very different from the regular classification one, and can be used by applications like targeted marketing, information retrieval, medical decision making, object recognition and intrusion detection. In fact, cost-sensitive classification can be used to express any finite-choice and bounded-loss supervised learning problems, and connects to popular machine learning problems including ranking, structured learning and online decision making.

Binary cost-sensitive classification problem considers only two kinds of costs: mis-predicting the first class as the second; mis-predicting the second class as the first. The studies on the problem can be traced back to the works of Elkan (2001) and Zadrozny (2003), which embeds the costs into the learning procedure by the technique of re-weighting the importance of each example. Multiclass cost-sensitive classification problem, on the other hand, can be more difficult than the binary one. There are several families of approaches:

1. by properly considering the costs during decision making rather than training, which tackles the problem from the Bayesian perspective (Domingos, 1999)
2. by re-weighting, which extends from the approach in binary classification (Zhou, 2006)
3. by reducing to regular classification, which generally needs re-weighting and re-labeling the examples (Abe, 2004; Lin, 2008)
4. by reducing to binary classification based on some different decomposition structures (Lin, 2008; Baygelzimer, 2005; Langford, 2005; Beygelzimer, 2007)

5. by reducing to regression by embedding the costs in the labels (Tu and Lin, 2010)

In addition to the algorithmic developments discussed above, cost-sensitive classification is being used in more and more applications in recent years. Selected applications that the speaker have first-hand experience include improving the performance in a real-world bacteria classification system (Jan, 2012), improving the ranking performance for information retrieval (Ruan, 2013), and constructing a useful model for recommender systems (Chen, 2011) which is part of the winning solution of the National Taiwan University team in KDD Cup 2011. The applications stimulate new needs for cost-sensitive classification, such as using the costs to approximate the true evaluation criteria of interest (Ruan, 2013), or being both cost-sensitive and error-sensitive (Jan, 2012).

In this tutorial, we review existing approaches of cost-sensitive classification, for both binary classification and multi-class classification. The approaches range from the Bayesian perspective of including costs during decision making to the reduction-based approaches that transform the cost-sensitive classification task to regular classification or regression tasks. We discuss the theoretical guarantees behind the approaches as well as their practical uses. We will also introduce some recent advances of cost-sensitive classification, including its success in applications. The tutorial assumes the basic background (techniques in classification and regression) in machine learning, and nothing more.

Bio:

Prof. Hsuan-Tien Lin received a B.S. in Computer Science and Information Engineering from National Taiwan University in 2001, an M.S. and a Ph.D. in Computer Science from California Institute of Technology in 2005 and 2008, respectively. He joined the Department of Computer Science and Information Engineering at National Taiwan University as an assistant professor in 2008, and has been an associate professor since August 2012.

Prof. Lin received the Distinguished Teaching Award from the university in 2011, and the Outstanding Mentoring Award from the university in 2013. He co-authored the introductory machine learning textbook Learning from Data. His research interests include theoretical foundations of machine learning, studies on new learning problems, and improvements on learning algorithms. He received the 2012 K.-T. Li Young Researcher Award from the ACM Taipei Chapter, and the 2013 D.-Y. Wu Memorial Award from National Science Council of Taiwan. He co-led the teams that won the third place of KDDCup 2009 slow track, the champion of KDDCup 2010, the double-champion of the two tracks in KDDCup 2011, the champion of track 2 in KDDCup 2012, and the double-champion of the two tracks in KDDCup 2013. He is currently serving as the Secretary General of Taiwanese Association for Artificial Intelligence.

Title: Recent Advances in Bayesian Methods

Presenter: Jun Zhu

Abstract:

This year marks the 250th Anniversary of Bayes theorem, which is playing an increasingly important role in statistical applications. Existing Bayesian models, especially nonparametric Bayesian methods, rely heavily on specially conceived priors to incorporate domain knowledge for discovering improved latent representations. While priors can affect posterior distributions through Bayes' theorem, recent work has shown that imposing posterior regularization is arguably more direct and in some cases can be more natural and easier. This tutorial will consist of two parts. First, I will review the recent developments of parametric and nonparametric Bayesian methods, with examples of Gaussian processes for regression, Dirichlet processes for clustering, and Indian buffet processes for latent feature learning. In the second part, I will introduce the connections between Bayes theorem and the principle of relative entropy minimization. In particular, I will introduce regularized Bayesian inference (RegBayes), a computational framework to perform posterior inference with regularization on the desired post-data posterior distributions. When the convex regularization is induced from a linear operator on the posterior distributions, RegBayes can be solved with convex analysis theory. Furthermore, I will present some concrete examples, including MedLDA for learning discriminative topic representations; infinite latent support vector machines for learning discriminative latent features for classification; and others on social network analysis, matrix factorization, multi-task learning, etc. All these models explore the large-margin idea in combination with a (nonparametric) Bayesian model for discovering predictive latent representations. I will discuss both variational and Monte Carlo methods for inference.

Bio:

Dr. Jun Zhu is an associate professor in the Department of Computer Science and Technology at Tsinghua University. His research interests lie in the development of statistical machine learning methods for solving scientific and engineering problems arising from artificial and biological learning, reasoning, and decision-making in the high-dimensional and dynamic worlds. Prof. Zhu received his Ph.D. in Computer Science from Tsinghua University. He did post-doctoral research in the Machine Learning Department at Carnegie Mellon University. His current work involves both the foundations of statistical learning, including theory and algorithms for probabilistic latent variable models, Bayesian nonparametrics, and large-margin learning; and the application of statistical learning in social network analysis, data mining, and multi-media data analysis. He is serving as an Area Chair at NIPS 2013 and a Local Chair at ICML 2014. He received Microsoft Research Fellow Award in 2007, China Computer Federation (CCF) Distinguished Dissertation Award in 2009, and the IEEE Intelligent Systems AIs 10 to Watch Award in 2013.

Workshop

Theory and Practice of Machine Learning

Organisers: Peter Sunehag, Marcus Hutter, Mark Reid

In this ACML 2013 workshop we intend to discuss what makes a theory relevant for practical development of algorithms and how the gap between theory and practice can be decreased. Several different forms of theoretical foundations for machine learning have been developed over the years for different settings. Conversely, there has been a recent and dramatic uptake in the application of machine learning techniques to an increasingly diverse range of problems. Arguably, some of the most successful algorithms on practical problems are not completely understood theoretically. We are interested in contributions relating to understanding or closing this gap, both specific technical contributions and general well-argued positions.

Conference Organisation

Organising Committee

General Co-chairs:
Wray Buntine (NICTA and ANU) and Bob Williamson (ANU and NICTA)

Program Co-chairs:
Cheng Soon Ong (NICTA) and Tu Bao Ho (JAIST, Japan)

Tutorial and Workshop Chair:
Lexing Xie (ANU)

Publication Chair:
Justin Domke (NICTA)

Web Chair:
Xinhua Zhang (NICTA)

Local Organisers:
Sarika Abbott (NICTA) and Krystel Boland (NICTA)

Steering Committee

- Tom Dietterich (Oregon State University, USA)
- Tu Bao Ho, Co-chair (JAIST, Japan)
- Bernhard Pfahringer (Waikato University, New Zealand)
- Masashi Sugiyama (Tokyo Institute of Technology, Japan)
- Takashi Washio (Osaka University, Japan)
- Geoff Webb (Monash University, Australia)
- Qiang Yang (Hong Kong University of Science and Technology, Hong Kong)
- Zhi-Hua Zhou, Co-chair (Nanjing University, China)

Senior Program Committee

- Aapo Hyvarinen (University of Helsinki)
- Baogang Hu (Chinese Academy of Sciences)
- Bernhard Pfahringer (University of Waikato)
- Chun-Nan Hsu (University of Southern California)
- Dell Zhang (Birbeck College University of London)
- Hang Li (Huawei Technologies)
- Hoai An Le Thi (University of Metz)
- Irwin King (Chinese University of Hong Kong)
- James Kwok (Hong Kong University of Science and Technology)
- Jieping Ye (Arizone State University)
- Kee-Eung Kim (KAIST)
- Kenji Fukumizu (Institute of Mathematical Sciences Tokyo)
- Kristian Kersting (University of Bonn)
- Marc Deisenroth (TU Darmstadt)
- Marco Cuturi (Kyoto University)
- Masashi Sugiyama (Tokyo Institute of Technology)
- Min-Ling Zhang (Southeast University)
- Pavel Laskov (University of Tuebingen)

- Ping Li (Cornell University)
- Scott Sanner (Australian National University)
- Shirish Shevade (Indian Institute of Science)
- Steven Hoi (NTU Singapore)
- Truyen Tran (Deakin University)
- Volker Roth (University of Basel)
- Xaiofei He (Zhejiang University)
- Yuh-Jye Lee (NTUST Taiwan)

Reviewers

- Ehsan Abbasnejad (NICTA)
- Marta Arias (Universitat Politecnica de Catalunya)
- Albert Au Yeung (Huawei Noah's Lab, Hong Kong)
- Albert Bifet (Yahoo! Research Barcelona)
- Battista Biggio (University ov Cagliari)
- Mathieu Blondel (Kobe University)
- Changyou Chen (ANU & NICTA)
- Hai-Leong Chieu (DSO National Laboratories)
- Jaedeug Choi (KAIST)
- Justin Domke (NICTA)
- Dinesh Garg (IBM IRL India)
- Roman Garnett (University of Bonn)
- Ricard Gavalda (Universitat Polit cnica de Catalunya)
- Shengbo Guo (Samsung Media Solution Center America)
- Sunil Gupta (Deakin University)
- Fabian Hadiji (University of Bonn)
- Toby Hocking (Tokyo Institute of Technology)
- Steven Hoi (NTU Singapore)

- Kaizhu Huang (Xi'an Jiaotong-Liverpool University)
- Baogang Hu (Chinese Academy of Sciences)
- Dino Ienco (Irstea Institute Montpellier France)
- Shiro Ikeda (Institute of Statistical Mathematics)
- Kazushi Ikeda (Nara Institute of Science and Technology)
- Adam Jatowt (Kyoto University)
- Zhuolin Jiang (Huawei Noah's Ark Lab)
- Takafumi Kanamori (Nagoya University)
- Masayuki Karasuyama (Kyoto University)
- Hisashi Kashima (University of Tokyo)
- Yoshinobu Kawahara (Osaka University)
- Marius Kloft (New York University)
- fabien lauer (University of Lorraine, Fance)
- Yuh-Jye Lee (NTUST Taiwan)
- Tam Le (Kyoto University)
- Hoai Minh Le (University of Lorraine, France)
- Xuelong Li (Chinese Academy of Sciences)
- Xiaol Li (Institute for Infocomm Research)
- Yu-Feng Li (Nanjing University)
- Bin Li (Nanyang Technological University)
- Shou-de Lin (National Taiwan University)
- Hsuan-Tien Lin (National Taiwan University, Taiwan)
- Zhiyong Liu (Chinese Academy of Sciences)
- Wei Liu (IBM T. J. Watson Research)
- Zhengdong Lu (Huawei Noah's Ark Lab)
- Hao Ma (Microsoft, USA)
- Tomoko Matsui (Institute of Statistical Mathematics)
- Daichi Mochihashi (Institute of Statistical Mathematics)

- Shinichi Nakajima (Nikon Corporation)
- Gerhard Neumann (TU Darmstadt)
- Marion Neumann (University of Bonn)
- Gang Niu (Tokyo Institute of Technology)
- Yung-Kyun Noh (Seoul National University)
- Sinno Pan (Institute for Infocomm Research, Singapore)
- Hsing-Kuo Pao (National Taiwan University of Science & Technology)
- Duc-Son Pham (Curtin University)
- Dinh Phung (Deakin University)
- Novi Quadrianto (University of Cambridge)
- John Quinn (Makerere University)
- Jan Ramon (K.U. Leuven)
- Matt Robards (NICTA)
- Benjamin Rubinstein (IBM Research)
- Sundararajan Sellamanickam (Microsoft Research INDIA)
- Anshumali Shrivastava (Cornell University)
- Yangqiu Song (HKUST)
- Mahita Sugiyama (Japan)
- Masashi Sugiyama (Tokyo Institute of Technology)
- Peter Sunehag (Australian National University)
- Taiji Suzuki (University of Tokyo)
- Akiko Takeda (University of Tokyo)
- Ichiro Takeuchi (Nagoya Institute of Technology)
- Dacheng Tao (University of Technology Sydney)
- Ryota Tomioka (University of Tokyo)
- Koji Tsuda (National Institute of Advanced Industrial Science and Technology, Japan)
- Tsuyoshi Ueno (ERATO Minato Project JST)
- Mirwaes Wahabzada (Fraunhofer IAIS)

- Jie Wang (Arizona State University)
- Lei Wang (Australian National University)
- Jingdong Wang (Microsoft Research Asia)
- Wei Wang (Nanjing University)
- Yu-Chiang Frank Wang (Research Center for Information Technology Innovation (CITI))
- Lei Wu (GE Global Research)
- Wei Wu (Microsoft Research Asia)
- Shuo Xiang (Arizona State University)
- Makoto Yamada (Yahoo! Labs)
- Haiqin Yang (Chinese University of Hong Kong)
- Haixuan Yang (Galway University, UK)
- Yi-Ren Yeh (National Taiwan University of Science and Technology)
- Yiming Ying (University of Exeter)
- Xiaotong Yuan (Department of Statistics and Biostatistics, Rutgers University)
- Lei Yuan (Dow AgroSciences)
- Junsong Yuan (Nanyang Technological University)
- Yu Zhang (Hong Kong Baptist University, China)
- Xinhua Zhang (NICTA)
- Changshui Zhang (Tsinghua University, China)
- Peilin Zhao (Nanyang Technological University)
- Tingting Zhao (Tokyo Institute of Technology)
- Jiayu Zhou (Arizona State University)
- Jinfeng Zhuang (Microsoft)
- Jun Zhu (Tsinghua University)
- Xingquan Zhu (University of Technology Sydney)
- Jianke Zhu (Zhejiang University)

Best Paper Award Selection Committee

- Hiroshi Motoda (Chair; Osaka University)
- Geoff Webb (Monash University)
- Masashi Sugiyama (Tokyo Institute of Technology)

Sponsors

We would like to thank the Air Force Office of Scientific Research, Asian Office of Aerospace Research and Development, NICTA, RuleQuest Research, and the Machine Learning Journal for their support.