EDITORIAL

Introduction: special issue of selected papers of ACML 2012

Zhi-Hua Zhou · Wee Sun Lee · Steven C.H. Hoi · Wray Buntine · Hiroshi Motoda

The fourth Asian Conference on Machine Learning (ACML 2012) was held in Singapore, during November 4–6, 2012. It follows the successful preceding ACML conferences held in Nanjing (China), Tokyo (Japan), and Taoyuan (Taiwan), in 2009, 2010 and 2011, respectively. This series of conferences aims to provide a leading international forum for researchers in machine learning and related fields to share their new ideas and achievements, and has attracted high quality submissions from Asia, Australasia, Europe and America.

ACML 2012 is the first machine learning conference to have executed a multiple-cycle submission procedure. There were 138 submissions received, where 99 were submitted for the first cycle and 39 were submitted for the second cycle. All submissions went through a selective review process. Every submission was double-blind reviewed, initially by at least

Editors: Zhi-Hua Zhou, Wee Sun Lee, Steven Hoi, Wray Buntine, and Hiroshi Motoda.

Z.-H. Zhou (✉)
National Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210023, China
e-mail: zhouzh@lamda.nju.edu.cn

W.S. Lee
Department of Computer Science, National University of Singapore, 117417 Singapore, Singapore
e-mail: leews@comp.nus.edu.sg

S.C.H. Hoi
School of Computer Engineering, Nanyang Technological University, Singapore 639798, Singapore
e-mail: chhoi@ntu.edu.sg

W. Buntine
National ICT Australia (NICTA), Locked Bag 8001, Canberra, ACT 2601, Australia
e-mail: Wray.Buntine@nicta.com.au

H. Motoda
Asian Office of Aerospace Research and Development (AOARD), Air Force Office of Scientific Research (AFOSR), Air Force Research Laboratory, USA, 7-23-17 Roppongi, Minato-ku, Tokyo 106-0032, Japan
e-mail: motoda@ar.sanken.osaka-u.ac.jp

4 reviewers from the program committee members according to their bidding results; every submission was also assigned to one primary and one secondary senior program committee member who oversaw the review process by moderating discussions among reviewers and resolving their divergences/conflicts. The Program Chairs and some additional reviewers examined borderline submissions to make confident decisions for those papers. Eventually, 12 papers were accepted for long oral presentation and 24 for short oral presentation. These accepted papers were published in the Journal of Machine Learning Research: Workshop and Conference Proceedings Series, vol. 25, 2012.

Authors of the best ten papers were invited to submit a significantly extended version to this special issue. The selection was made by the team of guest editors consisting of Program Chairs, General Chairs and Steering Committee Chairs, on the basis of the scientific quality and potential impact of these papers, as indicated by the conference reviewers. These ten extended papers have been reviewed again for two or three rounds of peer-review process according to the journal criteria, in a rigorous but efficient way, and for each paper some additional reviewers have been involved to enhance the journal review process. Finally, we present the following ten papers in this special issue.

The paper "Bayesian object matching", by Arto Klami, presents a full Bayesian approach to object matching in the context where no distance measure is available and the objects are represented by a feature vector. A latent binary permutation matrix for alignment is combined with a Bayesian Canonical Correlation Analysis as a basis for a distance measure. While the paper extends previous recent work, its approximate algorithms using both sampling and variational methods outperform previous algorithms on a variety of problems. Both a computational and an empirical analysis is presented.

The paper "Recovering networks from distance data", by Sandhya Prabhakaran, David Adametz, Karin J. Metzner, Alexander Böhm and Volker Roth, presents a probabilistic approach to reconstructing Gaussian graphical models by extending the central Wishart model to using a likelihood depending only on pairwise distances. This extension enables the model to be used with input of a kernel or distance matrix, without requiring direct access to the original vectors, and enables network reconstruction from any Mercer kernel, no matter whether the kernel is obtained from graph, string or more complex objects.

The paper "Correlated topographic analysis: Estimating an ordering of correlated components", by Hiroaki Sasaki, Michael U. Gutmann, Hayaru Shouno and Aapo Hyvärinen, presents the Correlated Topographic Analysis (CTA) method to estimate non-Gaussian components and their ordering (topology). It is assumed that components nearby on the topographic arrangement have both linear and energy correlations, whereas far away components are statistically independent. The ordering of components is fixed by considering these dependencies. The paper also presents a three-step optimization method which improves topographic estimation.

The paper "Variational Bayesian sparse additive matrix factorization", by Shinichi Nakajima, Masashi Sugiyama and S. Derin Babacan, investigates sparse matrix learning methods, and proposes a novel framework of sparse additive matrix factorization (SAMF). It systematically induces various types of sparsity by a Bayesian regularization effect, called model-induced regularization, which provides inference without requiring manual parameter tuning. An efficient iterative algorithm is proposed for the variational Bayesian approximation to SAMF, which provides a global optimal solution for a large subset of parameters in each step.

The paper "Conditional validity of inductive conformal predictors" by Vladimir Vovk, extends the theory of inductive conformal predictors (ICPs). Conformal prediction uses past experience to determine precise levels of confidence in new predictions. ICPs have better efficiency but by current theory have limited control of their coverage and are mostly de-

veloped for an online mode. This paper develops the theory of ICPs for the offline mode and extends the coverage to various forms of conditional probability. Examples are given including a demonstration on the Spambase data set.

The paper "Exploration and exploitation of scratch games", by Raphaël Féraud and Tanguy Urvoy, proposes Scratch games, a new variant of the multi-arm bandit problem. This setting is motivated by the problem of advertisement placement in online advertising, where a new advertisement may appear at any time and is used only for a known fixed number of times. Three algorithms, adapted from known methods, are proposed in the paper. Theoretical performance bounds that compare favorably with bounds for algorithms that do not take advantage of the properties of the problem are given for two of the algorithms, along with experimental results.

The paper "Hypervolume indicator and dominance reward based multi-objective Monte-Carlo tree search", by Weijia Wang and Michèle Sebag, investigates the use of Monte Carlo tree search algorithms which have been quite successful in numerous applications for multi-objective reinforcement learning. Two methods are proposed, one based on hypervolume indicator and the other based on Pareto dominance. Experiments show that the methods are able to handle problems with non-convex Pareto fronts and are competitive with a state-of-the-art method on an NP-hard grid scheduling problem.

The paper "Modeling individual email patterns over time with latent variable models", by Nicholas Navaroli, Christopher DuBois and Padhraic Smyth, investigates a latent variable modeling approach to extract information from individual email histories, focusing particularly on understanding how an individual communicates over time with recipients in their social network. The proposed model consists of latent groups of recipients, each of which is associated with a piecewise-constant Poisson rate overtime. Inference of group memberships, temporal change points, and rate parameters is carried out via Markov Chain Monte Carlo (MCMC) methods.

The paper "On using nearly-independent feature families for high precision and confidence", by Omid Madani, Manfred Georg and David Ross, investigates a late fusion problem and shows that combining the outputs of base classifiers trained on each feature type separately can substantially increase the recall of the combination at high precision when very different sources of evidence are available, This result is consistent with recent studies on multi-view learning. The authors also show how the probability of a joint false-positive mistake can be less than the product of individual probabilities of conditional false-positive mistakes.

The paper "Multi-stage classifier design", by Kirill Trapeznikov, Venkatesh Saligrama and David Castañón, investigates multi-stage classification systems where the training data are full, but during the testing phrase, measurements in a new modality can be acquired at each stage for an additional cost. They formulate a multi-stage reject classifier, where the classifier of each stage either classifies a sample using only the attributes acquired so far or rejects it to the next stage where more attributes can be acquired for a cost. They then solve the problem by developing an iterative algorithm in a boosting framework and present convergence and generalization results.