Machine Learning (2019) 108:717–719
https://doi.org/10.1007/s10994-019-05796-0

Foreword: special issue for the journal track of the 10th Asian Conference on Machine Learning (ACML 2018)

Masashi Sugiyama[1] · Yung-Kyun Noh[2]

Published online: 10 May 2019

We welcome you to this special issue of Machine Learning Journal (MLJ), comprising of papers accepted to the journal track of the 10th Asian Conference on Machine Learning (ACML 2018), held in Beijing, China, from 14 to 16 November 2018. The ACML conference is running a dedicated journal track alongside the usual conference proceedings track. We are delighted to share the contributions with you.

This year's ACML journal track received 35 submissions and 7 papers have been accepted for this special issue. The reviewing process was handled by the ACML journal track committee, consisting of Wray Buntine (Monash University, Australia), Steven Hoi (Singapore Management University, Singapore), Kee-Eung Kim (Korea Advanced Institute of Science and Technology, Korea), Wee Sun Lee (National University of Singapore, Singapore), Hang Li (Toutiao, China), Yu-Feng Li (Nanjing University, China) and Ivor Tsang (University of Technology Sydney, Australia), with Co-chairs and Editors of this special issue, Masashi Sugiyama (RIKEN/The University of Tokyo, Japan) and Yung-Kyun Noh (Seoul National University, Korea). Submitted papers were assigned to committee members by the Co-chairs according to the committee members' areas of expertise, while ensuring that there were no conflicts of interest. The committee members recommended reviewers and acted as meta-reviewers for the papers. Promising papers that did not quite meet the expected standard were allowed to be resubmitted after improvement, following the reviewing process of this journal.

The paper *Good Arm Identification* via *Bandit Feedback* by Hideaki Kano, Junya Honda, Kentaro Sakamaki, Kentaro Matsuura, Atsuyoshi Nakamura, and Masashi Sugiyama introduces a new concept of selecting a good arm which is not the best but useful. The Hybrid algorithm for the Dilemma of Confidence (HDoC) for this problem is proposed with theoretical bounds on the sample complexity. For many applications that not necessarily need the best choice, the proposed algorithm in this paper can be used for choosing good-enough choices after looking into a small number of samples.

✉ Yung-Kyun Noh
nohyung@hanyang.ac.kr

Masashi Sugiyama
sugi@k.u-tokyo.ac.jp

1 RIKEN/The University of Tokyo, Tokyo, Japan

2 Hanyang University, Seoul, Republic of Korea

The paper *Supervised Representation Learning for Multi-label Classification* by Ming Huang, Fuzhen Zhuang, Xiao Zhang, Xiang Ao, Zhengyu Niu, Min-Ling Zhang, and Qing He proposes a deep learning framework for multi-label learning by using a low-level unsupervised autoencoder and a high-level supervised softmax. The proposed setting allows users to learn the feature representations that are commonly used for all categories and to effectively reduce the necessary number of data. By using the proposed setting, multi-label classification performance is increased in the results of extensive experiments.

The paper *Bayesian Optimistic Kullback–Leibler Exploration* by Kanghoon Lee, Geon-Hyeong Kim, Pedro Ortega, Daniel D. Lee, and Kee-Eung Kim considers a Bayesian approach for the distribution of environment models, in order to significantly reduce the degree of exploration. The proposed algorithm, BOLKE, which a PAC-BAMDP algorithm, guarantees the near Bayes-optimality with high probability. The notion of bounded optimism is introduced for the proof of near Bayes-optimality and it generalizes the PAC-BAMDP proof steps. Comparisons with other well-known PAC-BAMDP methods, such as BOLT and BEB, are provided both theoretically and experimentally.

The paper *Annotation Cost-sensitive Active Learning by Tree Sampling* by Yu-Lin Tsou and Hsuan-Tien Lin considers a novel active learning situation for instances having different costs for annotation. For this situation, the cost-sensitive tree sampling (CSTS) algorithm is proposed for estimating the annotation cost and the utility at the same time. In order to simultaneously consider the cost and utility, tree structure is updated based on hierarchical sampling, and these two objectives are balanced by scaling the cost variance using the Gini impurity measure considering the worst case. The method is evaluated using both simulated and many real datasets.

The paper *N-ary Decomposition for Multi-class Classification* by Joey Tianyi Zhou, Ivor W. Tsang, Shen-shyang Ho, and Klaus-Robert Müller shows a new algorithm for solving multi-class classification. Instead of using the conventional binary-concept schemes such as one-versus-all or one-versus-one, the new algorithm considers meta-classes which are the combinations of several classes and decomposes the original multi-class problem into multi-class but simpler subproblems, using a divide-and-conquer strategy. A generalization error bound is derived which does not depend on the base classifier. The independence between subproblems makes the algorithm easily parallelizable.

The paper *Millionaire: A Hint-guided Approach for Crowdsourcing* by Bo Han, Quanming Yao, Yuangang Pan, Ivor W. Tsang, Xiaokui Xiao, Qiang Yang, and Masashi Sugiyama introduces a novel approach for collecting high-quality data in crowdsourcing. The proposed hybrid setting takes detour in the hint-guided stage when questions have uncertain answers. The mechanism is designed for high-quality workers to answer most questions directly for higher rewards, and the high-quality workers can be detected. As a result, high-quality labels can be collected. Incentive compatibility and the uniqueness of the proposed mechanism have been proven.

Finally, the paper *An Accelerated Variance Reducing Stochastic Method with Douglas–Rachford Splitting* by Jingchang Liu, Linli Xu, Shuheng Shen, and Qing Ling considers a proximal method for optimizing an empirical risk involving non-smoothness both in loss and regularizer. The proposed algorithm Prox2-SAGA, essentially performing the Douglas–Rachford splitting technique, generalizes the conventional Point-SAGA, which also uses a proximal operation on the loss terms but not on regularization. The theory in this paper provides convergence rates on both smooth and non-smooth losses as well as empirical experiments.

It has been a great honor for us to serve as Guest Editors for this special issue, and it would not have been possible without the contributions of many people. We wish to thank the journal

track committee for their efforts in ensuring the quality of this special issue, the referees for their time and effort in reviewing the papers and the authors for their contributions. We also would like to thank Peter Flach, Editor-in-Chief for MLJ, and Dragos Margineantu, Editor of special issues for MLJ, as well as the ACML Steering Committee for their guidance and support. Our gratitude also goes to Melissa Fearon and Karthika Deepak from the Springer editorial office for the help in ensuring that the process ran smoothly.