Zhang ML, Huang SJ, Long MS. Special section on learning from small samples. JOURNAL OF COMPUTER SCIENCE AND TECHNOLOGY 36(3): 588–589 May 2021. DOI 10.1007/s11390-021-0004-1

Preface

It is our great pleasure to announce the publication of this special section in JCST: "Learning from Small Samples".

Machine learning has achieved great success in various tasks. With the rapid growth of model size as in deep networks, the learning models become more and more complex, typically requiring a large scale of training samples with label annotations. However, in real-world applications, labeled data is usually limited. And it could be rather expensive to collect more labeled data because the labeling process is time consuming and requires domain expertise.

As a consequence, it becomes a major challenge to learn from a dataset with only a small amount of labeled samples. This special section is launched and aims to seek high-quality research work on learning from small samples. Each submission to this special section underwent a rigorous peer-review process overseen by the leading editor and the guest editors. At least two rounds of peer-review were carried out on each accepted paper. Finally, three contributions were selected for publication in this special section out of 26 submissions. These contributions show advanced technologies on learning from small samples, including disambiguation-based partial label learning, source-free domain adaptation, and deep cascade models.

The paper "Partial Label Learning via Conditional Label Aware Disambiguation" studies a weakly supervised learning problem, where each instance is partially annotated with a candidate set consisting of one ground-truth label and some other noisy labels. The authors propose to simultaneously train the model with label constraints and perform pseudo-labeling under a unified framework. The method can achieve state-of-the-art performance even when the training examples are not precisely labeled.

The paper "Source-Free Unsupervised Domain Adaptation with Sample Transport Learning" tries to learn from small samples by transferring knowledge from other related domains. The authors propose a source-free method to jointly model unsupervised domain adaptation and sample transport learning. The method does not require access to the source domain data, and thus could be widely applied to tasks with privacy concerns. Moreover, the method seeks the most valuable instances for labeling through an effective selection strategy, and can significantly reduce the number of labeled examples for training the model.

The paper "Multi-Scale Deep Cascade Bi-Forest for Electrocardiogram Biometric Recognition" considers an interesting application of electrocardiogram biometric recognition, where the key challenge is to obtain robust performance with small training data. The authors propose a deep cascade framework to incorporate multi-scale signal coding with deep cascade coding. By dealing with noisy signals with sparse representation and exploiting discriminative information with adaptive weighted pooling, the method achieves effective and robust results even with a small training set.

We would like to thank all authors for submitting their works to this special section and all reviewers for their great efforts in offering constructive and timely reviews on each submission. Last but not least, we are grateful to JCST for hosting this special section.

Leading Editor

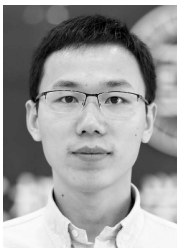
Min-Ling Zhang, Professor, School of Computer Science and Engineering, Southeast University, Nanjing
zhangml@seu.edu.cn

Guest Editors

Sheng-Jun Huang, Professor, School of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing huangsj@nuaa.edu.cn

Ming-Sheng Long, Associate Professor, School of Software, Tsinghua University, Beijing mingsheng@tsinghua.edu.cn

Min-Ling Zhang received his B.Sc., M.Sc., and Ph.D. degrees in computer science from Nanjing University, Nanjing, in 2001, 2004 and 2007, respectively. Currently, he is a professor at the School of Computer Science and Engineering, Southeast University, Nanjing. His main research interests include machine learning and data mining. In recent years, Dr. Zhang has served as the general co-chairs of ACML 2018, program co-chairs of PAKDD 2019, CCF-ICAI 2019, ACML 2017, CCFAI 2017, PRICAI 2016, senior PC member or area chair of AAAI 2017-2020, IJCAI 2017-2021, KDD 2021, ICDM 2015-2020, etc. He is also on the editorial board of ACM Transactions on Intelligent Systems and Technology, Neural Networks, Science China Information Sciences, Journal of Computer Science and Technology, Frontiers of Computer Science, etc.. Dr. Zhang is the steering committee member of ACML and PAKDD, secretary-general of the CAAI Machine Learning Society, and the standing committee member of the CCF Artificial Intelligence & Pattern Recognition Society. He is a distinguished member of CCF, CAAI, and a senior member of ACM and IEEE.



Sheng-Jun Huang is a professor in the College of Computer Science and Technology at Nanjing University of Aeronautics and Astronautics, Nanjing. He received his B.Sc. and Ph.D. degrees in computer science from Nanjing University, Nanjing, in 2008 and 2014, respectively. His research interests include machine learning and data mining. He has been selected to the Young Elite Scientists Sponsorship Program by CAST, and won the China Computer Federation Outstanding Doctoral Dissertation Award in 2015. He is a young associate editor of Frontiers of Computer Science and Journal of Computer Science and Technology.

Ming-Sheng Long received his B.E. degree in electrical engineering and Ph.D. degree in computer science, both from Tsinghua University, Beijing, in 2008 and 2014 respectively. He is an associate professor with the School of Software, Tsinghua University, Beijing. He was a visiting researcher with the AMPLab, UC Berkeley from 2014 to 2015. He serves as area chairs of major machine learning conferences (ICML/NeurIPS/ICLR). His research is dedicated to theories and algorithms of machine learning, with special interests in transfer learning, spacetime deep learning, and learning with rich knowledge.