Guest Editorial
Special Issue on Structured Multi-Output Learning: Modeling, Algorithm, Theory, and Applications

STRUCTURED multioutput learning is a topic in artificial intelligence that considers multiple structured outputs prediction for a given input. The output may involve structured objects in the form of sequence, string, tree, lattice, or graph and has values that are characterized by diverse data types, such as binary, nominal, ordinal, and real-valued variables. Such learning problems arise in a variety of real-world applications, ranging from document classification, computer emulation, sensor network analysis, concept-based information retrieval, and human action/causal induction to video analysis, image annotation/retrieval, gene function prediction, and brain science. As many complex real-world scenarios can be posed as a structured multioutput learning problem, their importance and popularity have been increasing steadily.

Within the machine-learning community, for example, structured multioutput learning has transpired in the form of multilabel/multiclass classification, multitarget regression, multiconcept retrieval (i.e., retrieval problems with query and documents containing multiple concepts), hierarchical classification with class taxonomies, label sequence learning, sequence alignment learning, supervised grammar learning, and many others.

In this special issue, the state-of-the-art scientific works in addressing some of the greatest challenges of structured multioutput learning are sought. In our call for article, we welcome all original submissions that focus on modeling, algorithm, theory, real-world applications, and related literature reviews/surveys of structured multioutput learning.

A total of 37 articles on structured multioutput learning have been received, out of which 15 high-quality articles have been chosen and accepted for publication in this special issue. The contents of these 15 articles are briefly discussed in what follows.

In the practical implementations of non-Gaussian statistical models, it is infeasible to arrive at an analytically tractable solution for estimating the posterior distributions of the parameters. Recently, the extended variational inference (EVI) with a number of non-Gaussian statistical models was considered in the article by Z. Ma *et al.* for multioutput learning. In their work, two core approximation strategies, namely, multiple and single lower bound approximations have been studied as possible analytically tractable solutions to the variational objective function.

Multioutput Gaussian processes (MOGPs) are extensions of classical Gaussian processes (GPs) for predicting multiple-output variables (also called channels/tasks) simultaneously. By modeling cross-channel dependences through cross convolution in the spectral domain, the article by K. Chen *et al.* proposed a new multioutput convolution spectral mixture (MOCSM) kernel for MOGPs in multioutput learning. MOCSM obtains the desirable property and can be reduced to the well-known spectral mixture (SM) kernel when a single channel is considered. The results of extensive experiments demonstrate the advantages of the proposed kernel and its promising performance in multioutput learning.

The interpretability of recurrent neural networks (RNNs) is a topic that has been less explored. To understand the inner mechanism of RNNs, B. J. Hou and Z. H. Zhou find that finite-state automaton (FSA) that processes sequential data offers greater interpretable inner mechanism according to the definition of interpretability and can be learned from RNNs as the interpretable structure. They proposed methods to learn FSA from RNN based on two different clustering methods. Moreover, they analyzed how the number of gates affects the performance of RNN and observe that the FSA learned from RNN gives semantic aggregated states. This article contributes to the theoretical understandings of RNN for multioutput learning.

Multilabel classification can significantly benefit from feature selection methods for handling high-dimensional sparse data set. However, the increasing complexity of multilabel feature selection, especially on continuous features, requires novel ways to manage data effectively and efficiently in distributed computing environments. To address this issue, J. Gonzlez-Lopez *et al.* presented a distributed model for mutual information (MI) adaptation involving continuous features and multiple labels on Apache Spark.

To better learn the correlations among image features, as well as labels, C. Li *et al.* presented a novel deep metric learning method to explore a latent space, where images and labels are embedded via two unique deep neural networks, respectively. They propose a two-way deep distance metric over the embedding space from two different views, i.e., the

Digital Object Identifier 10.1109/TNNLS.2020.2981981

distance between one image and its labels is not only smaller than those distances between the image and its labels' nearest neighbors but also smaller than the distances between the labels and other images corresponding to the labels' nearest neighbors. A new deep metric learning model for multioutput learning is proposed.

Label imbalance and noisy data that are commonly seen in the scenario of named entity recognition (NER) are largely neglected in the existing works. To address these two issues, J. Zhou *et al.* proposed a method termed robust sequence labeling (RoSeq) to handle the label imbalance issue by incorporating label statistics in the conditional random field (CRF) loss.

Large-scale multilabel learning (LMLL) assigns relevant labels for unseen data from a huge number of candidate labels. It is perceived that labels exhibit a long tail distribution in which a significant number of labels are tail labels. Most previous studies consider that the performance would benefit from incorporating tail labels. Nonetheless, it is not quantified how tail labels impact the performance. To address this issue, T. Wei and Y. Li showed that whatever labels are randomly missing or misclassified, the impact of labels on commonly used LMLL evaluation metrics (Propensity Score Precision (PSP)@k and Propensity Score nDCG (PSnDCG)@k) is directly related to the product of the label weights and label frequencies.

Network embedding is capable of providing low-dimensional feature representations for various machine-learning applications. H. Hong *et al.* taken advantage of network embedding in the view of multioutput learning. They proposed a generative adversarial network embedding (GANE) model to adapt the generative adversarial framework for better performance on multioutput problems.

Y. Zheng *et al.* developed a dynamically spatiotemporal regularization model for multioutput learning, which addresses the challenging boundary effects that exist in the correlation filter (CF)-based visual tracking.

Existing image subcategorization works relying on expert knowledge, and labeled images are both time-consuming and labor-intensive. To address these issues, Y. Yao *et al.* proposed to select and subsequently classify images into categories and subcategories. They first obtain a list of candidate subcategory labels from untagged corpora. Then, they purify these subcategory labels by calculating the relevance to the target category. This article proposed a novel approach to reduce the dependence on expert knowledge and labeled images in multioutput learning.

Zero-shot learning (ZSL), a type of structured multi-output learning, has attracted much attention due to its low requirement on training data in the target classes. Conventional ZSL methods usually project visual features into the semantic space and assign labels by finding their nearest prototypes. However, this type of nearest neighbor search (NNS)-based method often suffers from great performance degradation because of the nonuniform variances between different categories. H. Zhang *et al.* proposed a probabilistic framework by taking covariance into account to deal with the abovementioned issues.

Q. Li *et al.* proposed a caching strategy to replace the less frequently used kernel values in a novel SVM algorithm that adaptively achieves high hit ratios with little runtime overhead among different problems, including multilabel classification, multiclass classification, and multioutput regression problems.

Domain adaptation aims to exploit the supervision knowledge in a source domain for learning prediction models in a target domain. Z. Wang *et al.* proposed a novel representation learning-based domain adaptation method to transfer information from source to target domain where the labeled data are scarce in multioutput learning.

High-level semantic knowledge in addition to low-level visual cues is essentially crucial for cosaliency detection. Z. Zha *et al.* proposed a novel end-to-end deep learning approach for robust cosaliency detection by simultaneously learning high-level groupwise semantic representation as well as deep visual features of a given image group involving multioutput learning.

Last but not least, D. Xu *et al.* presented a survey to generalize different forms of multioutput learning in a single framework. It offers a comprehensive review and analysis of the multioutput learning paradigm. Some emerging challenges with multioutput learning from the perspective of the four Vs of Big data as potential research directions worthy of further studies have been delineated and discussed.

To summarize, the articles that appeared in this special issue reflect the state of the art in the fundamental research and development of multioutput learning and its applications to a wide range of relevant problems. Hence, the guest editors would like to thank all the authors for their utmost contributions in making this special issue a great success.

WEIWEI LIU, *Guest Editor*
School of Computer Science
Wuhan University
Wuhan 430072, China

XIAOBO SHEN, *Guest Editor*
School of Computer Science and Engineering
Nanjing University of Science and Technology
Nanjing 210093, China

YEW-SOON ONG, *Guest Editor*
School of Computer Science and Engineering
Nanyang Technological University
Singapore 637460

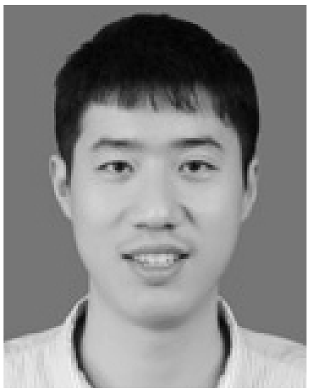
IVOR W. TSANG, *Guest Editor*
Centre for Artificial Intelligence
University of Technology Sydney
Ultimo, NSW 2007, Australia

CHEN GONG, *Guest Editor*
School of Computer Science and Engineering
Nanjing University of Science and Technology
Nanjing 210093, China

VLADIMIR PAVLOVIC, *Guest Editor*
Department of Computer Science
Rutgers University
New Brunswick, NJ 08901 USA

Weiwei Liu received the Ph.D. degree in computer science from the University of Technology Sydney, Ultimo, NSW, Australia, in 2017, under the supervision of Prof. I. W. Tsang.

He is currently a Full Professor with the School of Computer Science, Wuhan University, Wuhan, China. His research results have been published at prestigious journals and leading conferences, such as the *Journal of Machine Learning Research*, the IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, the IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, the IEEE TRANSACTIONS ON IMAGE PROCESSING, the IEEE TRANSACTIONS ON CYBERNETICS, NIPS, ICML, AAAI, and IJCAI. His current research interest includes statistical machine learning, which includes scalable multilabel classification, scalable multiclass learning, ultrahigh-dimensional feature selection, and large-scale clustering.



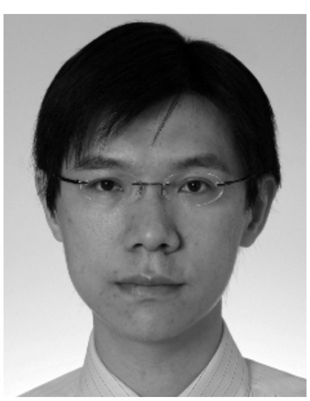
Xiaobo Shen received the B.Sc. and Ph.D. degrees from the School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing, China, in 2011 and 2017, respectively.

He is currently a Professor with the School of Computer Science and Engineering, Nanjing University of Science and Technology. He has authored over 30 technical papers in prominent journals and conferences, such as the IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, the IEEE TRANSACTIONS ON IMAGE PROCESSING, the IEEE TRANSACTIONS ON CYBERNETICS, NIPS, ACM MM, AAAI, and IJCAI. His primary research interests are multiview learning, multilabel learning, network embedding, and hashing.

Yew-Soon Ong received the Ph.D. degree in artificial intelligence in complex design from the Computational Engineering and Design Center, University of Southampton, Southampton, U.K., in 2003.

He is currently a Professor and the Chair of the School of Computer Science and Engineering, Nanyang Technological University, Singapore, where he is also the Director of the Data Science and Artificial Intelligence Research Center and a Principal Investigator of the Data Analytics and Complex System Programme, Rolls-Royce@NTU Corporate Laboratory, Singapore. His current research interests include computational intelligence span across memetic computation, complex design optimization, and machine learning.

Dr. Ong was a recipient of the 2015 IEEE Computational Intelligence Magazine Outstanding Paper Award and the 2012 IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION Outstanding Paper Award for his work pertaining to memetic computation. He is the Founding Editor-in-Chief of the IEEE TRANSACTIONS ON EMERGING TOPICS IN COMPUTATIONAL INTELLIGENCE and an Associate Editor of the IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION, the IEEE TRANSACTIONS ON NEURAL NETWORK AND LEARNING SYSTEMS, and the IEEE TRANSACTIONS ON CYBERNETICS.



Ivor W. Tsang is currently an ARC Future Fellow and a Professor of artificial intelligence with the University of Technology Sydney (UTS), Sydney, NSW, Australia, where he is also the Research Director of the UTS Flagship Research Centre for Artificial Intelligence. He has authored or coauthored over 180 research articles published in top-tier journal and conference papers. According to Google Scholar, his H-index is 55. His current research interests include transfer learning, feature selection, crowdsourcing intelligence, big data analytics for data with trillions of dimensions, and their applications to computer vision and pattern recognition.

Prof. Tsang was a recipient of the 2008 Natural Science Award (Class II) by Ministry of Education, China, in 2009, which recognized his contributions to kernel methods, the Australian Research Council Future Fellowship for his research regarding machine learning on big data in 2013, the 2019 International Consortium of Chinese Mathematicians Best Paper Award in recognition of his work "Towards ultrahigh dimensional feature selection for big data," which was published in the *Journal of Machine Learning Research*, the prestigious IEEE TRANSACTIONS ON NEURAL NETWORKS Outstanding 2004 Paper Award in 2007, the 2014 IEEE TRANSACTIONS ON MULTIMEDIA Prize Paper Award, and a number of best paper awards and honors from reputable international conferences, including the Best Student Paper Award at the IEEE Conference on Computer Vision and Pattern Recognition 2010.

Chen Gong received the B.E. degree from the East China University of Science and Technology (ECUST), Shanghai, China, in 2010, and the Ph.D. degree from Shanghai Jiao Tong University (SJTU), Shanghai, in 2016.

He is currently a Full Professor with the School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing, China. His research interests mainly include machine learning, data mining, and learning-based vision problems. He has published more than 30 technical articles at prominent journals and conferences, such as the IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, the IEEE TRANSACTIONS ON IMAGE PROCESSING, the IEEE TRANSACTIONS ON CYBERNETICS, CVPR, AAAI, IJCAI, and ICDM.

Dr. Gong also serves as a PC Member/Reviewer for several leading conferences, such as AAAI, IJCAI, ICDM, ACML, ACPR, and ICTAI. He received the National Scholarship by the Ministry of Education in 2013 and 2014, the Excellent Self-Financed Overseas Student Scholarship from the China Scholarship Council in 2015, the IBM Excellent Student Scholarship in 2015, the Excellent Doctoral Dissertation by Shanghai Jiao Tong University in 2016, and the Excellent Doctoral Dissertation by the Chinese Association for Artificial Intelligence in 2017. He has been a Reviewer for more than 20 prestigious journals, such as PR, the IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS (TNNLS), the IEEE TRANSACTIONS ON IMAGE PROCESSING (TIP), the IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING (TKDE), the IEEE TRANSACTIONS ON CYBERNETICS (TCYB), the IEEE TRANSACTIONS ON MULTIMEDIA (TMM), the IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS (TITS), and *ACM TKDD*.

Vladimir Pavlovic (Senior Member, IEEE) received the Ph.D. degree in electrical engineering from the University of Illinois at Urbana–Champaign, Champaign, IL, USA, in 1999.

He is currently a Professor with the Department of Computer Science, Rutgers University, New Brunswick, NJ, USA. From 1999 to 2001, he was a Member of Research Staff with the Cambridge Research Laboratory, Cambridge, MA, USA. In 2002, he was a Research Professor with the Bioinformatics Program, Boston University, Boston, MA, USA. His research interests include probabilistic system modeling, time-series analysis, computer vision, and bioinformatics.