Editorial

Advanced learning for large-scale heterogeneous computing

Coming with the era of big data, nowadays we have witnessed a drastic growth of heterogeneous data information involved in scalable computing and analysis. It has opened up a new gate and aroused new challenges in developing effective machine learning technologies to cater for such a heterogeneous computing requirement. Exemplar scenarios include, but not limited to, cross-modality visual search, machine translation, multi-modality medical imaging, and analyzing heterogeneous generic features. Under this circumstance, hybrid data comes from multiple sources, and is also typically hybrid in different feature channels. This requires a specific treatment covering different stages like data acquisition, storage, filtering, knowledge discovery, and classifier training. Among almost all stages above, principled machine learning techniques are indeed strongly expected for intelligent knowledge representations and smart decisions. In the latest years, we have encountered the wave of a variety of machine learning algorithms and frameworks, ranging from deep learning to quantum learning machines. It therefore becomes vital to report the very recent progress in advanced machine learning methodologies and state-of-the-arts for handling large-scale heterogeneous data. Our special issue is targeting at inviting original research works in this field, covering new theories, new algorithms, new implementations, new benchmarks, and new industrial deployments concentrating on the topic, **Large-Scale Heterogeneous Computing**.

Submissions came from an open call for paper and with the assistance of professional referees, 16 papers are finally selected out from in total 28 submissions after rigorous reviews. These papers focused on the Large-Scale Heterogeneous machine learning methods, and applied these methods to computer vision, bioinformatics and optimization. According to the applications, we categorize three subtopics for our special issue.

The first part contains 4 papers that are related to computer vision. The first paper, "Quad Binary Pattern and Its Application in Mean-Shift Tracking", proposed a new local texture descriptor—Quad Binary Pattern (QBP) and showed its effectiveness on target tracking by developing and incorporating a joint color-QBP model into mean-shift tracking. The second paper, "Combining Multi-representation for Multimedia Event Detection Using Co-training", introduced a novel Multimedia Event Detection system using multi-view learning, which integrated CNNs feature and Fisher Vector encoding of SIFT. In the third paper "Person Re-identification Based on Multi-Instance Multi-Label Learning", Lin et al. introduced an adapted canopy-k-means algorithm to evaluate the discriminability of image patches for pedestrian feature representation. And they firstly proposed a framework of attributes detection in pedestrian re-identification by applying Multi-Instance Multi-Label Learning (MIMLL). In the fourth paper "Recognizing Spontaneous Micro-Expression from Eye Region", Duan et al. presented a framework to recognize micro-expressions from eye region, which validated that the eye region is more discriminative than the whole face as well as mouth region for happy and disgust micro-expressions.

The second part contains 6 papers on bioinformatics application. The paper "Exploring Local Discriminative Information from Evolutionary Profiles for Cytokine-Receptor Interaction Prediction" proposed a machine learning based method for cytokine-receptor interaction prediction, which sufficiently captured local discriminative information from evolutionary profiles to represent true or non-true cytokine-receptor interactions and made predictions by Random Forest classifier. In the paper "iEnhancer-PsedeKNC: identification of enhancers and their subgroups based on Pseudo degenerate kmer nucleotide composition", Liu introduced a new computational predictor for enhancer identification, which employed a two level framework and incorporated the sequence-order effects. In the paper "Prediction of protein N-formylation and comparison with N-acetylation based on a feature selection method", Zou et al. developed a computational approach to identify N-formylated methionines. They also compared N-formylation with N-acetylation, another type of N-terminal modification of methionines. Their results furtherly helped the understanding of the two types of modifications and provided guidance for related validation experiments. In the paper "Identification of novel proliferative diabetic retinopathy related genes on protein-protein interaction network", Zhang et al. identified some novel candidate genes related to proliferative diabetic retinopathy using a protein-protein interaction network with a shortest path approach. In the paper" mGOF-Loc: A Novel Ensemble Learning Method for Human Protein Subcellular Localization Prediction", Wei et al. proposed a novel ensemble learning method for human protein subcellular localization prediction, which extracted a comprehensive feature set from different views to represent query proteins and then used a new multi-label ensemble classifier to make predictions. In the paper "Learning from Real Imbalanced Data of 14-3-3 Proteins Binding Specificity", Li et al. describes a novel method for predicting the binding affinities of 14-3-3 proteins. Several new features and machine learning methods are incorporated into the proposed predictor, which would be a useful computational tool in this field.

The third part contains 6 papers and focuses on the optimization techniques and machine learning. In the paper "Parametric and Nonparametric Residual Vector Quantization Optimizations

for ANN Search", Guo et al. introduced two optimized solutions based on residual vector quantization, which have good superiority on restoring quantization error. In the paper "An Improved Gravitational Search Algorithm for Green Partner Selection in Virtual Enterprises", Xiao et al. introduced carbon emission into the partner selection problem, and proposed an improved algorithm I-GSA/PSO that combined gravitational search algorithm and particle swarm optimization to solve G-PSP in virtual enterprises. In the paper "Dynamic Programming based Optimized Product Quantization for Approximate Nearest Neighbor Search", Cai et al. introduced an improved product quantization for approximate nearest neighbor search, which employed dynamic programming to optimize the subspace quantizers. In the paper "Expert Listwise Ranking method based on the Sparse Learning", Wang et al. introduced an expert listwise Learning to Rank method based on sparse learning, Which combined ranking loss function with feature dimension selection of sparse learning to reduce feature dimension. It can optimize ranking feature and improve the accuracy of ranking effectively. In the paper "Investment Behavior Prediction in Heterogeneous Information Network", Zeng et al. built an investment behavior prediction model of meta-path-based heterogeneous network, which considered multiple entity and relation types associated with the investment behavior of a particular investor. In the paper "Terminal neural computing: finite-time convergence and the related applications", a recurrent neural network named terminal neural network is developed and analyzed for solving time-varying matrix inversion. In addition, the TNN method is applied to trajectory planning of redundant manipulator. Both theoretical analysis and simulation results validate the effectiveness of the proposed neural network.

To conclude, papers in this special issue cover several emerging topics of advanced learning techniques and applications for large-scale heterogeneous computing. We highly hope this special issue can attract concentrated attentions in the related fields. We would like to thank Prof. Tom Heskes for providing the chance to organize this special issue. We also thank the reviewers for their efforts to guarantee the high quality of this special issue. Finally, we thank all the authors who have contributed to this special issue.

Quan Zou[*]
Tianjin University, China
E-mail address: zouquan@nclab.net

Wei Liu
IBM T.J.Watson Research Center, USA

Michele Merler
IBM T.J.Watson Research Center, USA

Rongrong Ji
Xiamen University, China

* Corresponding author.