



Special issue on semantic data analytics and bioinformatics

Haiying Wang¹ · Man-Wai Mak² · Hui Wang¹

Published online: 25 November 2017

© Springer-Verlag GmbH Germany, part of Springer Nature 2017

This special issue aims to report recent progresses in the areas spanning computer science, web technology, computational biology and Bioinformatics. It is intended to serve as a forum for researchers to share their research findings and their insights. After the rigorous two-round review process and stringent evaluation, a total of eight papers have been selected to be included in this special issue.

The first paper in this special issue focuses on named entity recognition (NER), which is one of the key tasks in biomedical text mining. Based on reviewing the three biomedical NER models and two well-known machine learning methods, Wang and his colleagues systematically compared six biomedical NER tools and their performance using two widely used corpora, GENETAG and JNLPBA. A multiplex representation of images was proposed in the second paper by Wei et al. They evaluated the representation scheme experimentally in face recognition and demonstrated its great potential for image analysis.

Gene Ontology (GO) is becoming the de facto standard for annotating gene products. However, its significance is not limited to annotation applications. Milano et al. investigated the information content (IC) associated with each GO term and how changes of annotations affect IC, which can be used for deriving functional similarity between gene products. The study conducted by Browne et al. demonstrates the feasibility of applying the knowledge encoded in GO to functional genomic research. GO-based similarities

and enrichment analysis have been incorporated into their integrative network-driven framework to uncover potential Alzheimer disease (AD) candidate genes.

Due to the ability to learn from data, machine learning has attracted increasing attentions in the area of computational biology and bioinformatics. This special issue features four papers in this field. Guo et al. proposed a method based on integrated support vector machines (SVM) with a hybrid kernel to predict protein interaction sites which is crucial for the understanding of the mechanism of protein–protein interaction and their cellular functions. Based on the incorporation of an adaptive-decision scheme into multi-label SVM classifiers, Wan and Mak introduced a new approach for predicting subcellular localization of multi-location proteins, which outperforms existing state-of-the-art multi-location predictors. In an attempt to extract subsets of coherently expressed genes and conditions from time-series gene expression data, Xue and her colleagues developed two bi-clustering algorithms which have potential to identify bi-clusters with statistical significance and biological relevance.

The special issue is a result of the dedication of many contributors. We are especially grateful to the anonymous reviewers for devoting their time and expertise which guaranteed the quality level achieved in this special issue. We would like to thank the Editors of the *International Journal of Machine Learning and Cybernetics* for their support and encouragement.

✉ Haiying Wang
hy.wang@ulster.ac.uk

Man-Wai Mak
man.wai.mak@polyu.edu.hk

Hui Wang
h.wang@ulster.ac.uk

¹ School of Computing, Ulster University, Northern Ireland, UK

² The Hong Kong Polytechnic University, Hung Hom, Hong Kong