

Editorial

Artificial Intelligence and Data Mining: Algorithms and Applications

Jianhong (Cecilia) Xia,¹ Fuding Xie,² Yong Zhang,³ and Craig Caulfield⁴

¹ Curtin University, Perth, WA, Australia

² School of Urban and Environmental Sciences, Liaoning Normal University, Dalian, China

³ School of Computer and Information Technology, Liaoning Normal University, Dalian, China

⁴ Edith Cowan University, Perth, WA, Australia

Correspondence should be addressed to Jianhong (Cecilia) Xia; c.xia@curtin.edu.au

Received 20 June 2013; Accepted 20 June 2013

Copyright © 2013 Jianhong (Cecilia) Xia et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Artificial intelligence and data mining techniques have been used in many domains to solve classification, segmentation, association, diagnosis, and prediction problems. The overall aim of this special issue is to open a discussion among researchers actively working on algorithms and applications. The issue covers a wide variety of problems for computational intelligence, machine learning, time series analysis, remote sensing image mining, and pattern recognition. After a rigorous peer review process, 20 papers have been selected from 38 submissions. The accepted papers in this issue addressed the following topics: (i) advanced artificial intelligence and data mining techniques; (ii) computational intelligence in dynamic and uncertain environments; (iii) machine learning on massive datasets; (iv) time series data analysis; (v) Spatial data mining; algorithms and applications.

Among them, there are six papers on new algorithm model design and optimisation. In “*Dictionary learning based on nonnegative matrix factorization using parallel coordinate descent*” by Z. Tang et al., the authors propose a novel method for learning a nonnegative, overcomplete dictionary for sparse representation of nonnegative signals. In “*A cost-sensitive ensemble method for class-imbalanced datasets*” by Y. Zhang and D. Wang, a cost-sensitive ensemble method is developed to solve imbalanced data classification. The proposed method is based on cost-sensitive support vector machine (SVM) and query by committee (QBC). In “*Vision target tracker based on incremental dictionary learning and global and local classification*” by Y. Yang et al., a robust

global and local classification algorithm for visual target tracking in uncertain environment is suggested based on sparse representation. In “*Analysis of similarity/dissimilarity of DNA sequences based on chaos game representation*” by W. Deng and Y. Luan, the authors construct three kinds of CGR spaces and describe a DNA sequence by CGR-walk model. As an application, the authors compare the similarity/dissimilarity of exon-1 of β -globin genes for nine species. In “*A real-valued negative selection algorithm based on grid for anomaly detection*” by R. Zhang et al., a GB-RNSA algorithm is proposed for anomaly detection. In “*An enhanced Wu-Huberman algorithm with pole point selection strategy*” by Y. Sun and S. Ding, a novel pole point selection strategy for the Wu-Huberman algorithm is developed to filter pole points by introducing a sparse rate.

Several authors deal with different aspects of time series analysis. In “*Piecewise trend approximation: a ratio-based time series representation*” by J. Dan et al., a time series representation PTA is developed to improve the efficiency of time series data mining in high dimensional large databases. In “*A dynamic fuzzy cluster algorithm for time series*” by M. Ji et al., a dynamic fuzzy cluster (DFC) is proposed based on improved a Fuzzy C-Means (FCM) algorithm and key points. The proposed algorithm works by determining those time series whose class labels are vague and further partitions them into different clusters over time. In “*A new strategy for short-term load forecasting*” by Y. Yang et al., a hybrid model based on the seasonal ARIMA model and BP

neural network is presented to improve the short-term load forecasting accuracy.

Papers collected in this special issue also focus on spatial data mining: algorithms and applications. In “*Ecological vulnerability assessment integrating the spatial analysis technology with algorithms: a case of the wood-grass ecotone of Northeast China*” by Z. Qiao et al., an assessment model of ecological vulnerability is developed using the analytical hierarchy process and a spatial analysis method. In “*Algorithms and applications in grass growth monitoring*” by J. Liu et al., a double logistic function-fitting algorithm is used to retrieve phenophases for grasslands in Northern China from a consistently processed Moderate Resolution Imaging Spectroradiometer (MODIS) dataset, and the accuracy of the satellite-based estimates is assessed using field phenology observations. Results show that the proposed method is valid for accurately identifying vegetation phenology. In “*The sustainable island development evaluation model and its application based on the nonstructural decision fuzzy set*” by Q. Wang et al., the authors discuss and establish a sustainable development indicator system and model and adopt an entropy method and the nonstructural decision fuzzy set theoretical model to determine the weight of the evaluating indicators. In “*Spatiotemporal simulation of tourist town growth based on the cellular automata model: the case of Sanpo town in Hebei province*” by J. Yang et al., the authors use a tourism urbanization growth model to simulate and predict the spatiotemporal growth of Sanpo town in Hebei province. In “*Model for the assessment of seawater environmental quality based on multiobjective variable fuzzy set theory*” by L. Ke and H. Zhou, a model based on a multiobjective variable fuzzy set theory is presented to evaluate seawater environmental quality. In “*Seismic design value evaluation based on checking records and site geological conditions using artificial neural networks*” by T. Kerh et al., several improved computational neural network models are proposed to evaluate seismic design values based on checking records and site geological conditions.

Finally, other applied problems are also considered. For example, in “*Crude oil price prediction based on a dynamic correcting support vector regression machine*” by L. Shu-rong and G. Yu-lei, a new accurate method of predicting crude oil prices is presented, which is based on an ε -support vector regression (ε -SVR) machine with a dynamic correction factor overcoming forecasting errors. The authors also propose a hybrid RNA genetic algorithm (HRGA) with the position displacement idea of bare bones particle swarm optimization (PSO) changing the mutation operator. In “*Mathematical model based on BP neural network algorithm for the deflection identification of storage tank and calibration of tank capacity chart*” by C. Li et al., the proposed method has better performance in terms of tank capacity chart calibration accuracy compared with other existing approaches and has a strong practical significance. In “*A study on coastline extraction and its trend based on remote sensing image data mining*” by Y. Zhang et al., data mining theory is applied to the pretreatment of remote sensing images. In “*Land use patch generalization based on semantic priority*” by J. Yang et al., the authors establish a neighborhood analysis model

and patch features and simplify the narrow zones and the feature sidelines. In “*Nonstationary INAR(1) process with q th-order autocorrelation innovation*” by K. Yu et al., an integer-valued random walk process with q th-order autocorrelation is discussed.

Acknowledgments

The guest editors of this special issue would like to express their thanks to the authors who have submitted papers for consideration and the referees of the submitted papers.

Jianhong (Cecilia) Xia
Fuding Xie
Yong Zhang
Craig Caulfield