

Editorial

Dear Colleague:

Welcome to volume 18(2) of Intelligent Data Analysis Journal.

This issue of the IDA journal consists of ten articles, all representing various aspects of theoretical and applied research and development related to the field of Intelligent Data Analysis.

In the first article of this issue Chen discusses the concept of interactive visual cluster analysis and presents a study of star-coordinate based visualization models. The article proposes an optimized design that provides the most efficient approach to an interactive model visualization and validation. The article contains a set of experimental evaluation and demonstrates how such an approach could generate effective visualization along with minimizing the computational costs. Mishra *et al.* in the second article of this issue discuss clustering from a different perspective. The authors present an approach to web user clustering in which they integrate set and sequence similarity measures with a rough set algorithm. Their application and case study involves using a well-known news data set in which the authors present the quality of their clusters with a number of clustering algorithms in DBSCAN. Rougeb *et al.* in the next article of this issue propose a semi-non-parametric density estimation framework for analyzing high-dimensional data. Their approach is based on development of a generic framework that consists of three modules, a Bayesian Network learning module, a fuzzy clustering module and an Independent Component Analysis module. The approach is evaluated in a supervised classification problem.

Vallim *et al.* in the fourth article of this issue discuss the importance of change detection in data streams and propose a framework for an unsupervised and automatic change detection in data streams. The proposed framework consists of a density based clustering step followed by a novelty detection procedure that is based on entropy. This article reports on the performance of their proposed approach in a set of experiments where insights into how to design change detection in data streams are obtained. Yu *et al.* in the next article advocate that all imbalanced data analysis cases may not necessarily be harmful and propose an approach to pre-estimate the level of harmfulness, given a particular data imbalance situation. The article includes studies about the reasons for harmfulness due to class imbalance and propose a simple and ingenious strategy that involves using matrix based class separability to estimate the level of harmfulness due to class imbalance. The results show that under certain circumstances, a-priori information could guide us to achieve better results.

The sixth article of this issue by Kittiphattanabawon *et al.* is about association rule mining where the authors present a region-based method to selectively use different association measures to rank different regions in the data. For example, focusing on news relations, judged by the domain expert, it would be completely related, somehow related and unrelated. The approach presented in this article has been extensively evaluated using large amounts of news data where the results showed substantial advantage. The seventh article of this issue is also on association rule mining where Koh and Pears argue that most association rule mining methods only consider positive frequent item sets. They propose a complementary approach that also looks for negative association rules and in their search they only consider the presence or absence of itemsets that are strongly associated. Their proposed approach is compared with other methods where they demonstrate that their approach is able to generate set of rules which are more interesting.

Lee and Giraud-Carrier in the next article of this issue demonstrate the risk associated with using default implementations of learning algorithms, without knowing the consequences. Their study includes an implementation of a radial basis function and its integration with three data mining software packages where they show that the proposed approach performs like Naïve Bayes in most instances. With the experiments performed, the authors argue that their proposed approach can influence computational complexity, ensemble design and metalearning for algorithm selection. The article by Iqbal *et al.* is also about association rule mining with Bayesian networks in XML data. In this article the authors propose a model for identifying sensitive items related to sensitive XML association rules. The approach is based on using a combination of a Bayesian network-based measures along with a K2 algorithm that is used to generate Bayesian networks. Their proposed model is tested and compared in several case studies presented in their article where they demonstrate that their approach reduces the risks involved in XML association rule mining. The last article of this issue by Morales *et al.* is an application on satellite image analysis where the authors propose a novel method to identify transition regions in the same way that they are found in nature. Their experimental results presented in the paper demonstrates that their approach outperforms some commercial software.

In conclusion, with this issue of the IDA journal, which is Volume 18(2), we are glad to report continuous increase in submission of manuscripts to our journal for evaluation and publication and also requests for special issues. A special issue in planning is in the area of Business Analytics in Finance and Industry. The deadline for submission of papers for this special issue is March 15, 2014 (please see call for papers included in this issue of the IDA journal). Interested authors can contact Prof. Richard Weber, Department of Industrial Engineering, Universidad de Chile (rweber@dii.uchile.cl). We look forward to receiving your feedback along with more and more quality articles in both applied and theoretical research related to the field of IDA.

With our best wishes,
Dr. A. Famili
Editor-in-Chief