Editorial

Dear Colleague:

Welcome to volume 18(4) of Intelligent Data Analysis Journal.

This issue of the IDA journal consists of eleven articles, all covering various topics related to the theoretical and applied research in the field of Intelligent Data Analysis.

The first 4 articles of this issue are about clustering and unsupervised learning. In the first article, Novoselova discusses the cluster stability issue and propose a stability-based algorithm to estimate the individual clusters of the dendogram. The proposed algorithm is based on a repetitive construction of the hierarchy of clusters followed by the calculation of the original consensus matrix. The experiments reported in this article on two simulated data sets and additional comparative analysis on real data sets demonstrate the advantages of the proposed approach. Vahidipour *et al.* in the second article of this issue also discuss hierarchical clustering and consider using multiple methods to generate a set of dendograms. The authors compare several weighted combination methods where they use hierarchical clustering results to derive a consensus. Their results show that weighting combination of hierarchical clustering performs better than averaging. Hu *et al.* in the third article describe a framework for cluster improvements which involves user supervision at the feature level. The proposed method ranks all features based on the recent clusters using cluster based feature selection and presents a list of highly ranked features to the user for labeling. Their experimental results on several real data sets demonstrate the feature set obtained using the new interactive framework that can produce clusters that better match the user's expectations compared with unsupervised approach. Ebrahimi and Abdollahi in the fourth article of this group focus on privacy preserving clustering (PPC) and argue that most of the existing techniques are based on heuristic notions. They propose an ϵ-differential based algorithm to generate a perturbed data for PPC and use a wavelet transform approach for reduced number of dimensions and less noise in the data. In their experiments reported in this article they compare their approach with other algorithms based on utility and privacy guarantees.

The next three articles are on temporal data and frequent patterns. Dey *et al.* in the fifth article of this issue discuss the topic of neighborhood discovery and describe four algorithms that are appropriate for knowledge discovery in temporal data. Their approach is based on identifying temporal neighborhoods with distinct demarcations that are based on unequal depth discretization of the data. The article contains results on both synthetic and real data sets. The next article by Shaw and Gopalan is about discovering meaningful trajectories from dynamic objects. The idea proposed by the authors is to apply the association-based mining concepts in order to find the longest trajectories from the frequent trajectory patterns. The path in this case is derived by applying the modified apriori where frequent pattern tree methods are compared with a standard graph based methods. The approach given in this paper is also applicable to game theory. The seventh article of this issue by Liu is also on mining frequent patterns, and in this case is on uncertain univariate class of data. This is a class of data where attributes in the transactions are associated with a quantitative interval and a probability density function that assigns a probability to each value in the interval. The author proposes three algorithms that are different in terms of data format that the structures used to store the data. The experimental results presented in the paper

show that the proposed algorithms excel when applied to different uncertain data sets, such as air quality and weather conditions.

The last four articles of this issue are more along the line of applied research. Graovac introduces an approach for automated categorization of text documents which includes a new dissimilarity measure between profiles. The technique that is based on K nearest neighbors is evaluated on several document collections in different languages. The results presented confirm that although the technique is fairly simple, in many cases it produces better results than other n-gram based methods. Zhao *et al.* in the next article introduce a new loss function for classification that could be considered as an approximation of the classical 0–1 loss function. The proposed loss function is compared with several existing loss functions in classification of noisy data sets. The presented experimental results show that the proposed method works better on noisy data sets, noisy features and outliers. Vateekul *et al.* in the tenth article of this issue discuss the problem of hierarchical multi-label classification along with performance evaluation in this field. The authors propose a new induction system which is based on support vector machines. In their experiments presented in the article they evaluate their system performance focusing on gene function prediction. And finally in the last article of this issue, Fang *et al.* discuss the importance of integrative data analysis focusing on the domain of life sciences and propose an integrative gene selection that is based on filter method and association analysis. The authors apply association analysis to analyse microarray data simultaneously with biological knowledge. Their experiments are based on four cancer related data sets and two types of biological knowledge. Their results show that models derived from a combination of biological knowledge and microarray data outperform the ones that are just derived from biological data alone.

In conclusion, with this issue of the IDA journal, which is Volume 18(4), we are glad to report continuous increase in submission of manuscripts to our journal. We are also contacted by many researchers in this field who would like to assist us in the evaluation of manuscripts. In addition, this year's IDA symposium will be held from October 30th to November 1st, in Leuven, Belgium (http://www.ida2014.org/). 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