Intelligent Data Analysis 13 (2009) 189–190
DOI 10.3233/IDA-2009-0362
IOS Press


# Editorial

Dear Colleague:

Welcome to volume 13(2) of *Intelligent Data Analysis* – An international Journal.

The second issue of volume 13 consists of nine articles. These articles represent several topics in IDA covering both applied and theoretical research. They vary from data preprocessing to association rule mining, automatic algorithm selection and data analysis performance evaluation.

In the first article, Shaari et al. introduce a new method for outlier detection that is based on the rough sets theory. Their approach is based on a measurement index called Rough Set Outlier Factor Value which is used to identify outliers. Their studies include a set of experiments using ten benchmark data sets and comparison of their measurement index with some existing ones. Their results show better performance by the newly introduced index. Zhao and Liu, in the second article of this issue, discuss the importance of understanding feature interaction and proper feature selection and reduction in data analysis. They introduce a method for data structure that is specific for feature quality evaluation and propose an information-theoretic feature ranking mechanism that can identify feature interactions. Their article includes a set of experiments in which the newly introduced method is compared with some existing ones for some performance measures such as data structure, raking, time complexity and scalability in search for interacting features. Along the same line of research, Utkin in the third article of this issue, introduces a method based on Dempster-Shafer theory that is used for ranking objects in large data sets. The main feature of this method is that it allows one to deal with comparisons of arbitrary subsets of alternatives. A set of experiments for the evaluation of this approach are reported in this article.

In the fourth article of this issue, Peppa and Freitas discuss the issue of choosing the right algorithm for a given classification problem and propose an automatic approach to this, where a rule induction algorithm is tailored to a given application domain. Their approach is based on a grammar based genetic programming during which a required algorithm is constructed. The reported approach is evaluated on 5 data sets where the results show that the automatically constructed rule induction algorithm was competitive with human designed ones.

The next two articles are more on applied research. Rueda and Arciniegas apply Self-Organizing-Maps to identify meaningful associations in objects. The application domain is finance and the particular problem reported in this article is to identify the speculative attacks' real effects, focusing on a set of variables that are based on economic, financial, and legal structure of a financial system. Interesting results are reported in this article. The next article by Karli and Saygin is on analyzing spatio-temporal sequences using different time granularities where they propose two new techniques. Their main motivation is that current techniques for searching periodic patterns can only work for fine granularities. Their experiments show how efficiently their proposed approach can identify periodic spatio-temporal patterns at different time granularities. Ordonez in the next article investigates mathematical relationships between association rules and clustering and correlation where it is shown that sufficient statistics for clustering and correlation are the linear and quadratic sum of points, respectively. The article includes results of a number of experiments where clustering and correlation models are developed. The results

show that clustering models are more accurate when more clusters are generated while correlation models are faster to generate than clustering with less accuracy.

Yun in the eighth article of this issue discusses the problem of large number of patterns in item set mining and emphasize the need for an effective pruning of unimportant patterns. This article includes two new algorithms for mining weighted frequent itemsets. This approach, that involves using normalized weights within the weights range and according to the importance of items, generates fewer but important weighted frequent itemsets in large data bases with minimum support. The extensive performance evaluation study included in this article, shows how this algorithm outperforms some of the existing ones.

The last article of this issue is a review article by Bosnić and Kononenko where they discuss the need for accuracy estimates of individual predictions and give an overview of approaches to generate these estimates. The overview is based on three research fields, which are: perturbing learning data, use of unlabeled data and sensitivity analysis. The authors present two classes of reliability estimates.

In conclusion, we are now in our 13th year of publication. Since submission of quality articles has substantially increased and we have been approached my many colleagues for special issues, we are gradually increasing the number of articles per issue. We greatly appreciate the support and encouragements that we have received from the IDA community. We look forward to receiving more and more quality articles in both applied and theoretical research in the field of Intelligent Data Analysis.

With our best wishes,

Dr. A. Famili
*Editor-in-Chief*