

Intelligent Data Analysis 2 (1998) 163-164



Letter From the Editor

Dear Colleague:

Welcome to Volume 2(3) of Intelligent Data Analysis journal!

Intelligent Data Analysis (IDA) has become a major endeavor in many enterprises. Research in data analysis, whether performed by research institutes, industries or academia, has resulted in a variety of powerful tools and techniques. On the other hand, many technological developments have made it very easy and inexpensive to acquire and warehouse terabytes of data. However, in most cases, the real value of these data is not properly exploited as part of a daily business activity of these enterprises. This is due to many reasons:

- lack of knowledge about tools, techniques, expertise and the experience from other applications,
- improper use of tools, techniques, and methodologies,
- improper use of data, that is sometimes scattered in an enterprise,
- lack of knowledge about proper selection of steps to intelligently analyse the data.

One of the aims of this journal is to address the above issues, through efficient dissemination of vast amount of information that becomes available by the researchers and practitioners who are working in this field.

This issue of IDA journal consists of five articles and a conference overview. The first article by Z. Yuanhui et al. explains the application neural network and genetic algorithm methods to mine classification rules. Focusing on reducing the input features first, the authors apply their algorithm to a number of data sets for testing and evaluation. The results show that their method generates rules of better performance comparing with the decision tree approach and the number of extracted rules is less that decision tree algorithms such as C4.5. The second article, by P. Domingos proposes and evaluates a meta-learner that seeks to retain most of the accuracy gains of multiple model approaches, while generating a single comprehensible model. This method is based on reapplying the base learner to recover every form of useful knowledge and making it explicit. Results are shown to be promising.

A major task in data pre-processing is dimensionality reduction. The article by M. Partridge and R. Calvo introduces a fast and simple algorithm on how to approximately calculate the principle components of a data set so that its dimensions can be reduced. Their approach is tested and evaluated on two high-dimensional image data sets, one from a hand written digits and the other from a hand written character data bases. On both applications, the authors show a fast convergence rate, comparing with other methods introduced in the literature. The fourth article by S. Tsumoto addresses the problem of extracting rules which are plausible representation of domain 'experts' decision making process. By examining the characteristics of experts rules, Tsumoto proposes a new approach to extract plausible

1088-467X/98/\$19.00 © 1998 Elsevier Science B.V. All rights reserved. PII: S1088-467X(99)00007-4

rules. Evaluating his method on medical data, the author show that the induced rules correctly represent experts' decision processes.

Clustering in data analysis deals with grouping cases so that groups are very different from each other and whose members are very similar to each other. The last article, by E. Boudaillier and G. Hebrail, introduces a new interactive interface to helps users to interpret the results of a hierarchical clustering process. They report successful use of their system in clustering electric load curves and providing a nice user interface for a user to be interactively involved in the clustering process.

And finally, in this issue we have an article, by Frey et al. that gives an overview of ICML-97 International Conference on Machine Learning, held in Nashville TN. The article gives a general observation on the conference as a whole.

We encourage potential readers to provide feedback. We welcome short conference overviews, workshop summaries and lessons learned.

Best wishes,

A. Famili

Editor-in-Chief