Editorial

Dear Colleague:

Welcome to volume 7(3) of *Intelligent Data Analysis*!

This issue of IDA journal consists of five articles that are a mix of some very interesting topics in applied and theoretical research in the field of Intelligent Data Analysis.

Data mining tools and packages have been built for several years. The main question that comes to everyone's mind is that: what is the right tool for me? In the first article of this issue, Giraud-Carrier and Povel introduce a comprehensive framework for characterizing and evaluating data mining software tools. They focus on one had, on the business requirements of these tools, and on the other hand to assist users to select proper tools for their applications. The author's goal is not to address the algorithm capabilities of these tools and rather emphasize on what the users need for proper selection of a data mining software package. Llorⱥ and Garrel, in the second article, discuss the issue of reducing the storage requirements in instance-based learning algorithms and present a new approach to deal with this problem. Their approach is based on a reduced set of prototypes with evolutionary algorithms that has been implemented and evaluated extensively. The results have shown that the approach achieves better reduction rates in storage requirements without losses in accuracy. This approach is also suitable for attribute selection in data mining.

Many researchers have worked in search for useful patterns and explanations in the form of associations. Li, Ogihara and Shenghuo, in the next paper, propose a new similarity measure between basket datasets based on associations. This measure, which is founded on information entropy, finds a mapping between categorical attribute sets and it is demonstrated that it can be applied to mine distributed data sets. In the fourth article of this issue, Soonthornphisaj and Kijsirikul introduce a web page categorization algorithm and evaluate its performance with supervised learning algorithms, Co-Training and Expectation Maximization. Their results show that their algorithm was quite effective for web page categorization applications. In the last paper of this issue, Thornton discusses disparity issues in sensing tasks that involve production of a 'high level' signal from 'low level' signal sources. The article contains an abstraction theory, which helps to explain the nature of the problem and the way to its solution. The solution proposed is based on the use of supervised adaptive methods from artificial intelligence. The article contains the results from a set of experiments to evaluate the performance of this approach.

And finally, we would like to remind our readers about the fifth Intelligent Data Analysis symposium that will be held in Berlin, Germany from August 28–30, 2003. Interested researchers can refer to the IDA-2003 website at (www.ida2003.org) for more information. Like previous events, there will be a special issue of IDA journal dedicated to this conference, will be published in early 2004 and will contain 5–6 of the best papers of this symposium.

With best wishes,

A. Famili
Editor-in-Chief