

## Editorial

---

Dear Colleague:

Welcome to volume 9(5) of *Intelligent Data Analysis* – An International Journal!

The five articles in this issue of the IDA journal are related to some of the best research performed in two areas: (i) challenges in data analysis, and (ii) data quality. All five articles involve extensive applied research where the results of new methods and their evaluation are supported using various real world data. This has always been our primary objective to make sure that accepted articles contain extensive evaluation if new methods are introduced by the authors.

In the first article of this issue, Shen *et al.* investigate the need for clustering on collaborative data sets and develop models for collaborative environments. They introduce new collaborative clustering algorithms and evaluate their approach where they demonstrate that the new algorithms can converge to a local minimum at a fairly reasonable rate. Their experimental results demonstrate several advantages in collaborative clustering. Shenoy *et al.*, in the second article of this issue discuss data mining in dynamic environments. In particular, the authors of this research introduce association rule discovery in dynamic transaction data bases using Genetic Algorithms. Various forms of transactions and time complexities are considered by the authors where their study shows that their dynamic data mining algorithm outperforms some of the existing approaches in terms of execution time and scalability. This is without compromising the quality or completeness of the discovered rules.

The next three articles of this issue are about data quality. Nandi and Klawonn attempt to detect and identify ambiguities in the data where they introduce a modified mountain method. The proposed method also helps in identifying the best local approximation function in regression problems. The article includes their evaluation with a one dimensional problem with real data and its evaluation with a two dimensional problem using an artificial data set. Similarly, Agyemang *et al.* investigate WEB outlier detection and its importance in WEB data mining. The main motivation for this research was that ordinary outlier detection algorithms work with numeric data sets and cannot handle WEB data where typically contain multimedia. This research proposes a new algorithm for mining WEB content outliers and contains and evaluation of its performance.

The last article of this issue by Khoshgoftaar and Rebourts discusses the problem of noise in software measurement data and how a noise filter can improve the quality of training data sets in software quality projects. They present an ensemble filter that is a combination of partitioning and classification filters and introduce a unique measure for comparing the relative efficiencies. Their results include an empirical study focusing on various filters where they demonstrate that a conservative filtering approach using multiple learners can improve the efficiency of the filtering schemes.

And finally, the 6th International Symposium of Intelligent Data Analysis (IDA-2005) was held in Madrid from September 8–10. Details of conference information are still available at the IDA society home page at: <http://www.ida-society.org> or at the conference home page at: <http://www.ida-2005.org>. Plans are underway to publish a special issue of the IDA journal, in 2006. This special issue would include 5 or 6 extended versions of the best papers presented at this symposium. Over the last 10 years,

we have benefited from the participation of researchers and practitioners of the IDA field in this important bi-annual event.

With our best wishes,

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