

## Editorial

---

Dear Colleague:

Welcome to volume 9(6) of *Intelligent Data Analysis* – An International Journal!

The six articles in this issue of the of the IDA journal are related to: (i) classifiers and classification methods, and (ii) data quality and transformation. As usual, these articles involve extensive theoretical and applied research where the results of new methods and their evaluation are presented.

In the first article of this issue, Tsoumakas *et al.* present a new approach that stands between classifier selection and model fusion. The proposed approach consists of a new statistical measure for the selection of the best subgroup amongst different classification algorithms. It also includes subsequent fusion of the decision of the models in each subgroup with weighted voting methods. The results of this work show its worth investigating more into the development of simple heterogeneous ensemble methods such as evaluation and selection, voting and weighted voting instead of complex methods that require more input parameters, sometimes with questionable results. Papadimitriou and Terzidis, in the second article, present a novel framework for development fuzzy rule-based classification systems. Their approach is based on the adaptation of support vector algorithms that will be used to identify a support vector fuzzy inference system. This is expected to result in a robust learning system. The results of this research include four examples that vary from simple rules to fuzzy rules and their evaluation. Hu and Sung in the next article address the problem of consensus clustering that deals with combining multiple partitions of a set of objects into a set of single consolidated partition. The approach takes as input a set of cluster labelings without having access to the original data. It then introduces a distribution-based view of partitions and proposes a series of entropy-based distance functions for comparing various partitions. Consensus clustering essentially works based on the optimizations and given a candidate partition sets, it always searches for a centroid partition with the smallest distance to a set. The paper includes evaluation of this approach using both artificial and real data sets.

Li and Wong, in the fourth article, discuss the issue of describing and capturing significant differences between multiple classes of data and focus on emerging patterns to describe significant differences. Their argument is that a pattern may occur in one class of samples but not necessarily in the others. The approach proposed in this article uses the most general patterns to construct accurate classifiers. Using a medical application, their approach is evaluated to be superior to some linear classifiers such as C4.5 and comparable to non-linear classifiers such as SVM and KNN. Khoshgoftaar and Van Hulse, in the next article present a methodology for detecting noisy attributes in software measurement data sets. Their empirical study reported in this article consists of injecting simulated noise into one or more attributes of a data set that normally contains no class noise. Results of their study contain evaluating the effectiveness of their method based on domain expert's feedback. The last article of this issue by Hu *et al.*, discusses the problem of using Principle Component Analysis methods for quantifiable association mining and propose a new method to learn associations between itemsets in the form of "Ratio Rules". The approach includes a support measurement to weigh the importance of each rule with respect to the entire data set. Results of their evaluation on several data sets that demonstrate its performance for discovering latent association between itemsets in large data sets are given.

And finally, the 6th International Symposium of Intelligent Data Analysis (IDA-2005) was successfully held in Madrid from September 8–10. We are currently working on selecting 5–6 of the best papers from this conference to publish a special issue of the IDA journal, in 2006. The IDA council also selected Ljubljana- Slovenia as the site for the IDA-2007 Symposium. Details of this event will be available in a few months.

With our best wishes,

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