

Editorial

Dear Colleague:

Welcome to Volume 4 (3, 4) of *Intelligent Data Analysis*!

Since early 90s, we have witnessed an enormous effort to bring ideas and methodologies from Artificial Intelligence into the scope of data analysis. Academic research and industrial practice have generated an impressive amount of work, which has spanned in: (i) various algorithms for efficient and meaningful data preprocessing and data analysis, (ii) a large variety of paradigms, and (iii) testing and evaluation of these efforts in virtually every domain where they could be applied. Some of the successful examples are in finance, medical, manufacturing, aerospace, and more recently in electronic commerce and genomics.

This double issue of IDA consists of 13 articles. The first six articles present some of the best research and development work in the field of Intelligent Data Analysis. In the first article, El Imrani et al. investigate the use of genetic algorithms along with fuzzy clustering for multi-modal function optimization. Their research showed that without using any prior knowledge, this approach would result in better search performance than standard sharing techniques, especially in terms of quality of the solutions. In the second article, Savnik and Flach present a new approach to the discovery of multi-valued dependencies in relational databases. Their approach consists of implementation of two algorithms, a top-down and a bottom-up. They showed that the discovery of database dependencies can assist in discovery of useful patterns from relational databases. The third article by Kontkanen et al. is about visualization of high-dimensional data. They propose a data visualization scheme where similarity of two attribute vectors is determined indirectly by using a formal model of problem domain which is a Bayesian network model. The paper includes results from implementation and evaluation of this approach.

The following three articles are about methods for improving the knowledge discovery process by which a model of real word data is generated. Brijs et al. propose an integer programming model to solve the problem of selecting the most appropriate subset of characteristic rules from a rule discovery process. They demonstrate how the proposed model can control a user-defined level of overall quality of a discovered model in combination with a maximum reduction in the original rule set. Bouroumi et al., in the next article, propose a new approach for a two stage unsupervised pattern classification scheme. The two stages consist of an unsupervised fuzzy learning procedure and an optimization procedure involving the fuzzy c-means. Evaluation of this approach in artificial and real test data are given by the authors. The next article by Ho and Scott discusses the issue of attribute value grouping and contains a new approach to reduce decision tree fragmentation. They report that attribute value grouping can result in substantial reduction in the size of a decision tree and it would have no effect on the accuracy of the trees produced.

For the last seven articles, here follows a short summary of each article and its contribution to the field of Intelligent Data Analysis. Boutalis et al. present a new fuzzy algorithm for recognition of the K-nearest neighbours. The algorithm, that requires very few distance calculations, consists of two steps: finding K-nearest neighbours and classifying test patterns. The algorithm is suitable for many practical recognition problems with a fairly low misclassification rate. Hong et al., in the next article, propose a new method that combines rough-set theory and fuzzy-set theory to solve the problem of producing a set of maximally

certain and possible rules from quantitative data. Their method involves transforming each quantitative value into a fuzzy set of linguistic terms using membership functions and then calculates fuzzy lower and upper approximations. The proposed algorithm can solve conventional crisp-data problems by using degraded membership functions. Šprogar et al. discuss vector decision trees, that in contrary to common decision trees, can make more than just one suggestion per input sample and can express some relationships between decisions that were unknown before. They explain a software tool called DecRain for building vector decision trees and show some results and their comparisons to classical decision trees. Kalles et al., also have an article along the same line of research in which they propose modifications to the classical decision tree induction approach in order to handle set-valued attributes. They present the empirical consequences of using set-valued attributes, which results in a significant improvements in accuracy of induced trees. In the next article, McClean et al. discuss discovering rules from event histories. They propose a methodology for discretising cycles which utilises phase-type distributions where the phases correspond to the discretised classes. The article includes results using some real world data from a medical application. In the last two articles, first Bruha presents an overview of the knowledge discovery process and its methodology in the form of steps and explains as machine learning, pre-processing of data and post-processing of results. And finally, Lertpalangsunti and Chan introduce a framework for intelligent systems that is based on the concept of intercommunicating hybrids. The system, that has been implemented using a real-time expert system shell, allows use of neural network modules, rules and procedures encapsulated into blocks which can connect with each other as data flow diagrams from data processing. The article includes results from use of this system in a forecasting application.

And finally, we are very pleased with the level of support from the IDA community and encouraged by the number of quality papers submitted to our journal. We welcome both applied and theoretical papers related to the field of IDA. This double issue is an example of a mix of both types of articles that have been submitted and selected by our referees for our journal. Thank you.

Best wishes,

A. Famili
Editor-in-Chief