

Editorial

Dear Colleague:

Welcome to volume 13(5) of *Intelligent Data Analysis* – An international Journal.

This issue of volume 13 consists of eight articles that cover several areas, all related to the field of Intelligent Data Analysis. These consist of topics related to data preprocessing, modeling and evaluation and feature selection.

In the first article of this issue, Storari et al. discuss the importance of proper identification of association and correlation rules from data and propose two new methods for learning Bayesian networks from a given data set. Their approach is based on an improved K2 learning algorithm where association and correlation parameters are extensively exploited. Two modified versions of K2 algorithm, called K2-Lift and K2- X^2 are presented and evaluated in this paper where the results show better quality and learned network. Jensen and Shen in the next article discuss feature selection and propose use of a two level hierarchical nesting to calculate likelihood ratios when comparing two data sets of common origin. The main idea in this article is to select the best features without the need for expert knowledge. Results of several experiments are reported by the authors. The third article of this issue is by Cardoso and de Carvalho that emphasizes on cluster quality and evaluation of data clustering in an unsupervised learning task. Key issues related to clustering operation are discussed in this article and a general methodological approach is presented which includes the identification of appropriate thresholds for different cluster quality indices. Their proposed approach is compared with some of the methods listed in the literature.

Giraud-Carrier et al., in the fourth article discuss challenges in analyzing medical data where they introduce an approach to build predictive models for surgical procedure selection to achieve better surgical procedures. They suggest that observational studies supported by predictive informatics could complement classical clinical trials in medical research. Chan et al. investigate a new type of dynamic graph analysis where regions of a graph representing similar patterns are identified. This approach, which is a multi-level graph partitioning algorithm, could be useful in event detection and fault diagnosis, where regions are temporally and spatially separated. The proposed method is evaluated using one synthetic and two real data sets and the results are reported in this article.

The sixth article by Petrović et al. is about identification of the best features in textual data. Since efficient text mining requires feature selection and dimensionality reduction, the authors in this article report on using correspondence analysis as an alternative representation to better separate documents that belong to different classes. Results of their experimentation performed on a corpus of Croatian-English documents is reported. The next article by Madylova and Ögüdücü is also about grouping textual documents, where the authors investigate several similarity measures based on the semantic similarity of their terms. The objective of this research is to study the effects of semantic and single term similarity measures used in clustering documents (with the examples given in Turkish) where the results of their experiments on data from web sites are given. And finally, Pazouki and Rahmati in the last article of this issue present a framework for extracting information about particular players from a video stream. They present a set of data analysis techniques consisting of three stages that automatically extract certain features, such as player's speed, and traverse distance. The performance of their proposed system is reported in the article using some soccer video data.

In conclusion, we will hold the eighth International Symposium on IDA (<http://ida09.liris.cnrs.fr/>) this year in Lyon, France from August 31st to July 2nd. We are expecting to publish a special issue of IDA journal that will contain the extended version of some of the best papers presented in this conference. We are also expecting to have two more special issues from other events that were held in 2008-09. We look forward to receiving more and more quality articles in both applied and theoretical research in the field of Intelligent Data Analysis.

With our best wishes,

Dr. A. Famili
Editor-in-Chief